ESSAYS IN ECONOMICS OF ELECTRONIC COMMERCE

A Dissertation

by

ANIRBAN SENGUPTA

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

August 2007

Major Subject: Economics

ESSAYS IN ECONOMICS OF ELECTRONIC COMMERCE

A Dissertation

by

ANIRBAN SENGUPTA

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Approved by:

Chair of Committee,	Steven N. Wiggins
Committee Members,	Steven L. Puller
	Qi Li
	James Griffin
Head of the Department,	Amy J. Glass

August 2007

Major Subject: Economics

ABSTRACT

Essays in Economics of Electronic Commerce. (August 2007)

Anirban Sengupta, B.S., St. Xavier's College;

M.A., Jawaharlal Nehru University

Chair of Advisory Committee: Dr. Steven N. Wiggins

The advent of the internet has revolutionized the way people buy and sell. The internet is characterized by increased access to information. This increased information should foster convergence to the "law of one price," for homogenous goods.

The surge of electronic markets has motivated a stream of research focusing on comparing the efficiency of the internet market to the traditional one. This dissertation contributes to the existing literature of consumer search behavior in electronic markets and its effects on the price level and dispersion in the market more generally.

A part of this dissertation assesses the direct and indirect effect of increased internet usage on the prices of airline tickets, using a unique contemporaneous online and offline transaction data for airline tickets, covering the final quarter of 2004. The study also investigates the relationship between increased internet usage and price dispersion in the market for airline tickets. This study also includes an exhaustive set of controls for airline ticket characteristics, namely refundability, advance purchase requirements, travel and stay restrictions, class of travel, departure and return day of the

week and time, flight level load factor along with other market structure data used in the standard airlines literature.

The existing theoretical literature in consumer search extended to the electronic markets assumes, for simplicity, that all consumers in the internet markets are the "searchers," looking for the lowest price. The internet, however, also plays the role of a convenient shopping medium for a group of consumers whose primary motivation is not to search for the lowest price. The contemporary literature incorrectly categorizes these consumers as the traditional searchers. The remaining part of this dissertation provides a modification to the existing theoretical models of consumer search to accommodate both searchers and non-searchers in each of the electronic and traditional markets and derive the implications of the increased internet usage on the average level of prices and price dispersion in a market selling a homogenous good.

DEDICATION

This dissertation is dedicated to Ma and Baba. I also dedicate this work to my loving wife Ritu, whose love and support has been critical throughout this journey of graduate school. I also dedicate it to my Dada, who has been a perpetual source of encouragement to me. Finally, I dedicate it to my Aida and my late grandparents, whose love and affection I would always miss.

ACKNOWLEDGEMENTS

I would like to take this opportunity to thank my advisor Steven N. Wiggins for his enormous support and help in writing this dissertation. He has been a source of great inspiration to me. I owe him the success, whatever little, which I have achieved. He taught me the difference of being 'good' and being 'best'. I would like to thank him for the faith he reposed in me. I would also like to thank him for instilling the belief in myself.

I would like to thank Steve Puller who has been a constant source of support and help in these years. I would also like to thank my other committee member, James Griffin and Qi Li for their help.

I would also like to thank Manuelita Ureta for having helped me with some of the computational programming. This has been extremely critical for my research work. I would also like to take this opportunity to thank Rajiv Sarin and Hae-shin Hwang for having kept their trust on me in times of need.

Finally, I would like to thank my friend Birendra Rai who has been a source of huge support to me in the last several years. And finally, I would like to thank Tyffanne Rowan, who took care of me whenever I needed the most.

Thanks also to all my friends and colleagues and the department faculty and staff for making my time at Texas A&M University an experience of a lifetime.

TABLE OF CONTENTS

CHAPTER Page

LIST OF FIGURES

LIST OF TABLES

CHAPTER I

INTRODUCTION

The existence and the reasons of price dispersion have captivated the interests of the economists for a long time now. Price dispersion is the upshot of different sellers offering different prices for the same good in a given market. It differs from price discrimination where the sellers offer different prices to different groups of consumers or in different geographical areas.

There exists a varied theoretical literature that has tried to model price dispersion. The determinants of price dispersion are varied. Price dispersion can be attributed to both demand and supply side. The supply side drivers of price dispersion mainly constitute the number of firms, decreasing returns to scale and firm's marginal cost of production. On the consumer side, difference in demand elasticity among the consumers is an important determinant of price dispersion. Another important source of price dispersion is the imperfect information among consumers who are not equally informed about the lowest prices and the sellers offering them. This asymmetric information, resulting from the difference in the search behavior conditional on differential search cost, has been an important topic in the field of economics of information.

The relationship between consumer search behavior and price dispersion has been in the frontier of academic research in recent times. The renewed interest in

---

This dissertation follows the style of *The Journal of Industrial Economics*.

consumer search behavior and how it affects the average level of prices and price distribution can be attributed to the advent of the internet based electronic markets, which has redefined the ways in which goods are bought and sold.

An economically efficient market is one where a product is sold at the lowest possible price or at the marginal cost of production for a given quality. If all firms use the same technology and have the same cost of producing the given quality good a unique price is an equilibrium outcome. But in reality, different firms have asymmetric information about the price and quality of good. This generates a distribution of prices offered for the same good, rather than a single uniform price in the equilibrium. People looking for a better 'deal' have to search extensively to gather information on the different prices and products so that they can comparison shop and find the best 'deal' for themselves. This process has costs associated with it. The cost of search is any amount of money, time, or effort that a buyer may incur for obtaining price and quality information about the products. The failure to search adequately results in imperfect information among buyers while informational asymmetry among sellers provide some degree of market power so as to charge a price higher than the competitive price.

The advent of internet has revolutionized the way people buy and sell. This is fostered by improved access to information, particularly prices, products, quality of services and other product attributes. This improved information has made markets more competitive by exerting downward pressure on both level and dispersion in prices for goods bought and sold on the internet, and more generally.

Internet presumably reduces transaction costs thereby creating more efficient and frictionless markets as predicted by the ‘law of one price’ for homogenous goods. That is none, of the firms in the market should be able to sustain a price above the competitive level, resulting in uniform prices and eliminating any amount of dispersion that existed otherwise.

Consumer search behavior and their implications on both price levels and dispersion has been studied both theoretically and empirically. The maturity of electronic markets has motivated a stream of research with particular interest in comparing the efficiency between internet enabled electronic market and the traditional market.

The internet provides the opportunity to search more extensively and at a lower cost for particular goods and over numerous dimensions including price, product attributes and sellers. This improved access to information can be thought of as making consumers on the internet into “searchers”. Many models define searchers as those who have a low or zero search cost so that they search exhaustively (for example, they sample all the stores in a market). Such search enables them to buy at the lowest price. This definition fits well for internet consumers who by virtue of a mouse click get access to highly varied and exhaustive price information. Non-searchers in these models are defined as groups with substantial search cost prohibiting them to sample exhaustively. These individuals search till they find a price lower than their reservation price. Some models also characterize these non-searchers as convenience shoppers or who search for certain product attributes other than the price. This characteristic of sampling a limited number of stores characterizes consumers in the traditional markets well.

This dissertation extends the existing literature of information economics, both theoretically and empirically. The existing theoretical literature offers a reasonable analysis of the economic implications of consumer search on prices and price dispersion. Recent literature in electronic commerce has adhered to the traditional search models in addressing research pertaining to electronic markets: how average prices are affected as more people search on the internet, how increased internet usage affects price dispersion in different markets, to mention a few. Most of these models, however, do not characterize the electronic markets completely.

Internet serves a dual role. Firstly, it provides the consumers with a platform to search exhaustively and at a lower cost. Secondly, it serves as a channel for convenient mode of transaction for the more convenience shoppers. Most of the models fail to accommodate the convenience role of the internet. Existing research in electronic markets implicitly assume the internet consumers to be intrinsically shoppers, those looking for the lowest price. This assumption incorrectly categorizes those customers as shoppers who use the internet only as a convenient medium of transaction. Convenience of shopping on the internet mainly constitute of less transaction time, less effort and hassle free shopping experience. Additionally, the literature also assumes consumers in the traditional markets to be lesser price sensitive and categorize them to be non-searchers. More generally, there exist a group of consumers who searches extensively (at least, all stores in a region) to find the lowest price available or the ‘best deal', making them searchers as opposed to non-searchers in the more traditional markets.

Chapter IV of the dissertation presents a model that accommodates both groups of searchers and non-searchers in the electronic and traditional marketplace. I then use this modified market framework to derive implications of the increased access to internet on average prices and price dispersion in the overall market for a homogenous good, where the overall market is a linear combination of the online and traditional channels of sale.

The core of this dissertation consists of the analysis of the effects of the internet on airline pricing. The empirical analysis of the effects of electronic markets on prices and price dispersion has been of much interest to the empirical researchers in recent years. The focal point of these analyses has been primarily to compare market efficiencies between electronic and traditional markets. The theory of consumer search predicts prices should converge to competitive level and reduce dispersion as consumers become more aware and informed. The emergence of the electronic markets characterized by increased information has provided the platform for empirical testing of these different theoretical hypotheses.

The existing empirical literature has primarily focused on consumer durable goods like CDs, books, DVDs; financial products like term life insurance policies and automobiles. An integrated analysis on an industry basis assessing the role of the internet is still lacking, owing much to the lack of adequate data. Also, a substantial drawback to this existing empirical literature is the reliance on posted prices. Chapter

III of this dissertation provides the most comprehensive analysis to date of the effects of the internet on the pricing of any major industry as a whole by using a unique contemporaneous online and offline transaction data of airline ticket prices.

The travel industry in US is a multi-billion dollar industry. The market for airline tickets constitutes the core of the travel industry which also includes hotels, car rentals, cruises, tours and other services. The growth of internet usage for booking airline tickets has grown at a staggering rate since Alaska airlines sold the first ticket on the internet back in 1995. In 1996, sale of airline tickets through internet websites like Expedia or Travelocity made up for only 0.4 percent of the total sales. This share was estimated to be about 11 percent in 2003 and is increasing steadily.[1]

Despite this growing trend of internet booking for airline seats, economic analysis of how the internet has affected the airline ticket prices has been restricted. Much of this can be attributed to the lack of adequate data, particularly contemporaneous online and offline transaction data. Another drawback associated with the existing work in airline pricing in general, has been the inability to control for different ticket attributes like refundability, Saturday night stay-over, travel and stay restrictions, flight level load factor which significantly contributes to the variation in the ticket prices paid by the consumers. One challenge to isolate the effects of how the internet affects the airline ticket prices paid (if any) and to compare the online and offline prices, is to adequately

[1] Source: US Department of Transportation, July 2000 (http://www.oig.dot.gov/StreamFile?file=/data/pdfdocs/cr2000111.pdf)

control for the ticket characteristics. Chapter III of the dissertation specifically overcomes these limitations in existing literature.

Chapter III also provides one of the very few studies in the internet literature that uses contemporaneous online and offline transaction data for a specific period of time where the identity of the ultimate sellers is the same. The richness of the data provides a unique opportunity to test for the direct and indirect effects of the internet on domestic airline ticket prices in US.

Specifically, this chapter addresses three key issues pertaining to the direct and indirect effects of the internet. Firstly, it provides a direct comparison of the online and offline prices after controlling for all ticket characteristics, flight level load factors and market structure variables.

Models of consumer search predict that as more people search (on internet) the expected price to be paid by both searchers and non-searchers would decrease. In the spirit of this theoretical prediction, I estimate the indirect effects of internet, by measuring how increased internet usage affects the average ticket prices on a route, both for the internet and non-internet consumers.

The theoretical literature predicts a non-monotonic relationship between the fraction of searchers and price dispersion. Stahl (1989) argues that in a market with no searchers, monopoly price is the unique price equilibrium. As the share of searchers increases so does the dispersion in prices. In the limit, when all are searchers in the market competitive price is the equilibrium price with no price dispersion. To sum, moving from a market with no searchers to a market with all searchers price dispersion initially increases and then approaches zero. This study tests this theory by investigating the relationship between the fraction of people who buy on the internet (searchers) and price dispersion in the market (route) after controlling for all characteristics (ticket and market level). This study also revisits some established results in airlines using a more complete data. This study furthers our understanding of the dynamics of airline pricing along with the role of the internet, simultaneously filling the void in the existing literature.

The remainder of this dissertation is organized as follows. Chapter II provide an overview of the relevant literature in consumer search, internet and airlines. Chapter III provides an empirical study of the direct and indirect effects of the internet on prices for airline seats. Chapter II presents a new model of consumer search where we extend an existing sequential search model to cater to the dynamics of electronic markets. Finally, Chapter V presents the concluding remarks and scope of future research.

CHAPTER II

LITERATURE REVIEW

Overview

The advent of the Internet has revolutionized the way people buy and sell. Internet is characterized by increased access to information, particularly prices, products and quality of services. This increased information has considerably improved buyer information and consequently has put downward pressure on the level of prices and price dispersion in prices, both on the internet and more generally.

Increased information should foster convergence to the 'law of one price', for homogenous goods. That is, one would think, firms should not be able to sustain a price above the competitive levels, resulting in uniform prices and eliminating dispersion.

Consumer search behavior and their implications on price levels and dispersion has been studied both theoretically and empirically. The surge of electronic markets has motivated a stream of research on online prices and dispersion particularly focusing on comparing the efficiency of the Internet market to that of the traditional one.

The internet provides the opportunity to search more extensively and at a lower cost for particular goods and over numerous dimensions including price, product attributes and sellers. This improved access to information can be thought of as making consumers on the internet into "searchers". Many models define searchers as those who have a low or zero search cost so that they search exhaustively (for example, they sample all the stores in a market). Such search enables them to buy at the lowest price. This definition fits well for internet consumers who by virtue of a mouse click get access to highly varied and exhaustive price information. Non-searchers in these models are defined as groups with substantial search cost prohibiting them to sample exhaustively. These individuals search until they find a price lower than their reservation price. Some models also characterize these non-searchers as convenience shoppers or who search for certain product attributes other than the price. This characteristic of sampling a limited number of stores characterizes consumers in the traditional markets well.

Theory of Consumer Search: A Literature Review

Theoretical literature analyzing heterogeneity in consumer search behavior and its implications on price and price dispersion in extensive. The first economic model addressing consumers search behavior dates back to Stigler (1961). Since then the search models have undergone varied modification incorporating randomized price setting behavior by the firms, sequential search in oligopolistic markets to mention a few.

Price dispersion is the outcome of different sellers offering different prices for the same good in a given market (Hopkins (2006)). Price dispersion differs from price discrimination where the sellers offer different prices to different groups of consumers or in different geographical areas. This phenomenon of price dispersion can be attributed to both supply and demand side. The supply side drivers of price dispersion primarily constitute of the number of firms, decreasing returns to scale and firms marginal cost of production. The difference in demand elasticity among customers is one of the main drivers of dispersion on the demand side. Also, imperfect information among consumers is a key driver of price dispersion. The role of imperfect information arising due to heterogeneity in search costs among consumers has been of much significance with the maturity of the electronic markets in recent times.

Earlier models of price dispersion failed to explain dispersion as an equilibrium phenomenon. The need to develop such a theoretical framework was first emphasized by Rothschild (1974) who identified some serious drawbacks with the then existing models of price dispersion. Those models posited a ‘partial-partial’ equilibria focusing only on one side of the market, failing to provide analyses that would evaluate price dispersion as an equilibrium outcome. The challenge was thus to show that offering a range of prices would be an equilibrium response by sellers to the search behavior of consumers, and vice versa.

The shortcomings in the then existing models was highlighted by earlier work of Diamond (1971). He shows that with imperfect information among consumers monopoly pricing is the equilibrium outcome rather than price dispersion. Diamond (1971) uses a simple theoretical framework to derive this result. The model assumes a large number of identical consumers with unit demand at a price, p, offered by a large number of identical sellers. The consumers buy as long as the offered price is less than a reservation price, r. The model also assumes that the consumers know the distribution of prices but each knows the price of only one of the sellers in the market. Beyond the first price quote, each consumer faces a fixed cost of search for each additional price quote. In this set-up, the optimal search rule is to buy from the first seller if the price offered by the seller is not greater than the reservation price, r. In such a framework, the optimal price for the sellers is to charge the reservation price, r, with no price dispersion in the equilibrium. This equilibrium satisfies Rothschild's (1974) criteria where the consumer behavior is optimal as paying to learn additional prices is a waste of effort; monopoly price is optimal; as where there is no search, there is no incentive for the sellers to cut prices to increase sales.

Since then models of price dispersion (Salop and Stiglitz (1977), Varian (1980)) have departed from the traditional framework of Diamond (1971) by assuming away sequential search approach and adopting what has been come to known as 'clearing house models' (Baye and Morgan (2004)). The 'clearing house' literature largely follows Varian (1980). One key assumption of these models is that sellers face a constant marginal cost, c. In such a framework there exists no pure strategy equilibrium for the sellers as long as there remains a group of informed and uninformed consumers. The motivation is if all sellers charge the same price then there exist the incentive for an individual seller to deviate from this price and undercut to attract the informed buyers. On the other hand, charging the monopoly price to the positive mass of uninformed consumers assures the sellers a minimum profit. Finally, charging the reservation price, r, is the optimal strategy when the price of other sellers are close to c. In this framework, there exists a symmetric mixed equilibrium in which all sellers randomize according to the same continuous distribution of prices.

Varian (1980) assumes heterogeneity in search costs among buyers, where buyers facing a higher search cost choose to remain uninformed. Burdett and Judd (1983) shows that it is possible to close the clearinghouse model even if all consumers face the same cost to become informed (Edison (2003)). In such a framework, Burdett and Judd (1983) argue that people who choose to become informed pay the lowest price while those who choose to be uniformed would expect to pay the expected value of the distribution. Burdett and Judd (1983) derives the condition that for a sufficiently low search cost the equilibrium distribution of prices is sufficiently dispersed such as to make the consumers indifferent between paying or remaining informed.

A question that arises is whether heterogeneity in search costs leads to more robust price dispersed equilibria, either under sequential search or clearing house models. Consumers with a range of search costs are unlikely to search in absence of price variation, especially when the sellers charge the monopoly price. With sufficient seller heterogeneity, all sellers charging the monopoly price may not be a sustainable equilibrium. If one replaces unit demand function by a continuous one, cost heterogeneity among sellers can lead to heterogeneity in monopoly prices such that in equilibrium low cost sellers charge the monopoly price while high cost sellers charge a price lower than the monopoly price to make positive sales (Reinganum (1979)).

Theoretical models have also tried to rationalize the phenomenon of price dispersion associated with a homogenous good by invoking ex-ante consumer heterogeneity with respect to search costs (Shilony, 1977), information availability (Stiglitz and Salop, 1977), the propensity to search (Wilde and Schwartz, 1979) or ex-ante producer heterogeneity arising from differences in production costs (Reinganum, 1979), and marketing strategies (Varian, 1980). Burdett and Judd (1983) argue that the critical aspect of a model that leads to equilibrium price dispersion is that there must exist an ex-post heterogeneity in consumer information which can arise with or without ex-ante heterogeneities among the consumers and/or the producers.

Equilibrium price dispersion can probably be most easily rationalized by the presence of a positive marginal cost of obtaining each price quote. This serves as a plausible explanation of price dispersion in both conventional (Pratt et al (1979), Carlson and Pescatrice (1980), Sorensen (2000)) as well as for some virtual markets (Smith, Bailey and Brynjolfsson (1999), Bakos (2001)). The marginal cost of search in conventional markets involves visiting an additional store, while in the online market it involves visiting an additional website if no price comparison site exists, or visiting the website of the actual vendor if the price comparison site only involves the listing of the prices or `deals' of the different vendors on their webpage (example Kelkoo.com (Baye, Morgan and Scholten (2004)).

Stahl (1989) uses a sequential search model with heterogeneous search costs among consumers to rationalize price dispersion as an equilibrium outcome. Stahl (1989) assumes a market for a homogenous good with two groups of consumers, shoppers and non-shoppers. The firms offer a price that is drawn randomly from a given distribution. Using this framework, Stahl (1989) presents a non-monotonic relationship between the fraction of searchers and dispersion where the dispersion initially rises for low fraction of searchers but then monotonically decreases.

Presumably, the internet lowers the search cost for consumers. In this spirit, Stahl's model provides the appropriate theoretical framework to study price dispersion in the presence of an internet market for a homogeneous good.

To summarize, the existing models explaining the consumer search have some straightforward implications. Firstly, both consumer and social welfare are decreasing with search costs. A reduction in search cost for consumers create a positive externality by inducing more competition and decreasing the average level of prices, for both searchers and non-searchers. Further, increased search also reduces dispersion in prices. Finally, we have seen that models with homogenous sellers give rise to mixed equilibria while models with bilateral heterogeneity generate a pure equilibrium.

Empirical Analysis of Consumer Search: A Literature Review

Introduction of the internet revolutionized the way people search for goods. The improved search that the internet provides led to a surge of interest among the economists to study its impact on prices and price dispersion in different markets.

There has been considerable research in recent years that have attempted to investigate if internet creates a ‘frictionless’ market and converges prices to “law of one price”. Recent empirical research has also attempted to compare the distribution of prices in online and offline markets and see if more improved information in the online markets cause the dispersion to shrink. This section provides a comprehensive overview of the more important empirical research findings in recent years.

Empirical literature analyzing the effects of the internet on prices, both levels and dispersion, is mixed. Bakos (1997) was the first to empirically find evidence of lower dispersion in the internet markets as compared to the physical retailers, attributing it to the typically lower search costs in the internet markets. Brynjolfsson and Smith (2000) compared the prices and dispersion for online and offline markets individually, for twenty titles of books and compact discs (CDs) over a one year period to find similar amount of dispersion in the two markets. They found that on average online prices were lower while the dispersion in the two markets was comparable. The dispersion in the online markets was marginally lower once the market shares of the online retailers (measured by web traffic) were accounted for. Lee and Gosain (2002) reported similar findings for the market of music CDs.

Contrary to the above findings Erevelles, Rolland and Srinivasan (2001) finds persistent higher dispersion in the internet markets for vitamins when compared to other channels of physical retailing namely drug stores, discount retailers, super markets and warehouse retailers. They also find higher average prices in the electronic retailers as compared to the traditional retailers.

Clemons, Hann and Hitt (2001) studies dispersion in the airline ticket prices across different online travel agents. After controlling for different ticket attributes namely Saturday night stay-over, time of arrival and departure, number of connections or stop over they find prices vary by almost 18 percent across the different online websites. One argument for this apparently large dispersion can be attributed to the lack of control for other product heterogeneities namely ticket characteristics like refundability, advance purchase restriction, travel restrictions, meal offering among others, that have been found to explain the variations in the prices to a great degree.

The mixed bag evidences of comparisons of price dispersion and average prices between conventional and internet markets, has contributed to the ongoing research in this field. All of these earlier works, however, suffer from one crucial shortcoming. Most of these studies have been done in the nascent stage of maturity of electronic markets such that much of the higher dispersion in these markets can be attributed to the lack of maturity of the electronic marketplaces. The implications of the internet on levels and dispersion of prices can be better understood with analyses done in matured electronic markets.

Brown and Goolsbee (2002) investigated the effects of increased internet usage on the average prices and dispersion for term life insurance policies. Using data for almost 30,000 transactions purchased from 46 different term life insurance providers between 1992 and 1998, they found that for every ten percent increase in the share of people using the internet for search purpose decreased the average prices by about 8 percent. This paper, is one of the very few in the literature to use actual transaction data to investigate the effects of internet on average prices and level of dispersion, spanning a period from where the internet was non-existent to a period when the internet was being used quite significantly for search purpose.

Another study in a similar spirit is by Morton, Zettelmeyer and Silva-Rissio (2001) who studied the impact of internet car referral service on the prices charged by the car dealers in California, for the period between 1999 and 2000. They find that people who searched on the internet on average paid two percent less than those who did not. The study also found that as share of people who were referred by the website increased, the spread in the prices charged by the dealers decreased, suggesting that increased internet usage decreased the dispersion.

More recent studies of price dispersion show that the level of price dispersion in the internet market is decreasing over time. Pan et. all (2003) provided the most recent evidence on online price dispersion. Comparing prices for different products line books, CDs, DVDs, computers and other varieties of consumer electronic goods, covering a time period between November 2000 and February 2003; they report a 10 percent decline in average prices from 38.5 percent to about 28 percent. These findings are consistent to what Baye, Morgan and Scholten (2004) finds by studying online monthly prices for thirty six popular consumer electronic products listed on one of the established websites, Shopper.com. The findings suggest a substantial decrease in price difference from seventy to thirty percent in a period of eighteen months.

A variety of other factors can also be argued to contribute to the persistent dispersion in the internet markets. Brand loyalty (Lal and Sarvary (1999), bundling of products (Varian (1980)), difference in service (Pan et. all (2002)) quality are some of them. Analysis of these factors affecting the price dispersion lie beyond the scope of this dissertation and hence is left for future research.

Drawbacks in Existing Literature

The existing literature on consumer search behavior extended to include the electronic market suffers from some limitation, both theoretically as well as empirically.

The traditional consumer search cost models serve as backbones to more recent literature analyzing online and offline markets. For simplicity purposes, most of the e-commerce models implicitly assume the internet consumers to be the searchers and the consumers in the traditional markets to be non-searchers. This assumption though allowing for empirical tractability undermines the dual role served by the internet. On one hand, internet provides an opportunity for the consumers to search extensively and secure the 'best' deal for themselves. Additionally, the internet also allows for a convenient and time efficient channel of transaction for consumers who necessarily are not motivated to find the best 'deal'. The existing theoretical literature assumes away the group of consumers who use the internet for convenience, thereby excluding the group of low intensive searchers in the internet market.

Similarly, the assumption of all consumers being non-searchers or low intensive searchers in the offline market is a strong one. It is not hard to realize of a substantial fraction of consumers in the offline market, who would search extensively to find the lowest price or best 'deal' for themselves. A part of this dissertation builds on these motivations to include both searchers and non-searchers in each of the online and offline market, and derive implication for average prices and price dispersion in markets for homogenous good given such a market structure.

I build on the basic theoretical framework of Stahl (1989) to include the group of more and less informed consumers in each market and derive its implications for the average price and price dispersion in the overall market, where the overall market is a linear combination of the offline and online markets.

Stahl's (1989) model gathered increasing popularity among empirical economists, especially those studying the effects of the electronic markets on price levels and dispersion, primarily due to the ease of taking this model to real time data. All of these empirical studies assume that Internet presumably reduces search cost significantly and as a consequence (unless specific data is available) the internet (online) consumers are characterized as the searchers while those in the traditional markets are assumed to be typically non-searchers. Using this definition of consumer heterogeneity, many empirical studies have tried in the recent past to evaluate the effects of increased internet usage on both the levels and dispersion in prices in varied markets.

Such empirical work, barring a few, is often criticized on the basis of two arguments. Firstly, some amount of dispersion can be explained by product heterogeneity and not differences in search costs. A second line of skepticism is that dispersion in posted prices may not be consonant with uniformity in prices actually paid. Those who post higher prices may not sell anything such that it requires prices to be weighed by the market shares of the sellers resulting in less dispersed prices than if all sellers are given equal weights. A final line of criticism lie in the comparison between the posted 'national' internet prices with prices from disparate local stores. The local price, though from a traditional physical store, can be region specific and may not be representative of the national level of prices which the internet prices are on average. In this spirit, prices recorded from the local stores (Scholten and Smith (2001)) may not be comparable to the posted prices on the internet.

Contributions to the Existing Literature

Chapter III of this dissertation contributes to the existing literature of both electronic and airlines market. This study analyzes the direct and indirect effects of the internet usage on airline ticket prices. This study also simultaneously overcomes the limitations present in both literatures. This study uses contemporaneous online and offline transaction data of airline tickets. The study also uses information on individual ticket characteristics that provides unique control over product heterogeneity that earlier studies in airlines and in electronic markets could not. Further, the data also allows to compare the effects of the internet on airline prices in a ‘well defined market’, where a market is a specific route and hence do not suffer from the criticism of comparing ‘national’ internet prices with local regional offline prices, as the earlier studies. This study, allows furthering our understanding of the dynamics of the internet and how it affects the prices and dispersion more comprehensively and accurately than what the existing literature has to offer.

Chapter III of this dissertation contributes to the literature of both e-commerce and airlines industry by providing the most comprehensive analysis of the effects of the internet on the prices for airline tickets, both directly and indirectly. This study also uses a very rare contemporaneous online and offline transaction data for airline tickets to provide an integrated analysis of the internet's effects on prices for one of the more important industries, airlines.

I use this unique transaction data to addresses three important issues in reference to the airline industry and electronic markets. Firstly, this chapter quantifies the price differential for airline tickets between the electronic and traditional markets. Secondly, it investigates how the increased use of Internet affects the average prices for airline tickets, for both online and offline consumers. Finally, it measures how increased internet usage contributes to the price dispersion in the market for airline tickets. All of these issues are investigated while controlling for an exhaustive set of controls that influences the variation in ticket prices, namely refundability, advance purchase requirement, travel restrictions, minimum and maximum stay restrictions, load factor at the time of purchase and other set of hedonic factors which the existing literature analyzing airline pricing has not been able to do. Compared to the existing literature, this study provides the most comprehensive analysis of airline ticket prices, to date.

The theoretical literature on consumer search behavior segregates the consumers based on their propensity to search. This framework has been extended to the literature when comparing the Internet and traditional markets, as is done in Chapter III. The internet however, serves a dual role. Firstly, internet is used by many as an efficient medium of comprehensive search at a lower cost. On the other hand, many use the internet for convenience purposes, not necessarily for finding the lowest price or the best deal. In this spirit, consumers in the electronic markets can be categorized as both searchers (who search extensively for the lowest price or the best deal on the Internet) and non-searchers (those who use the Internet only for convenience). This aspect of the Internet has not been included in the theoretical literature of search, especially in reference to the Internet markets.

Chapter IV, of this dissertation, undertakes this theoretical exercise to include a group of searchers and non-searchers in each of the electronic and traditional markets individually and to derive the implications of the increased internet usage on the average level of prices and price dispersion, in a market selling a homogenous good.

In sum, this dissertation addresses the issues of the consumer search behavior in the internet markets and how it affects the level and dispersion of prices. This dissertation contributes to the economic literature by providing a comprehensive analysis of the effects of consumer search behavior in the electronic markets and more generally. It also provides a detailed analysis of the effects of the internet on the prices and price dispersion in the airline tickets market, which the existing literature does not offer. It also proposes a theoretical framework for future empirical research to accommodate the searchers and non-searchers in the both traditional and electronic markets. This compilation of research would extend our understanding of both the electronic and the traditional markets along with the economic implications of consumer search behavior on how these markets behave.

CHAPTER III

AIRLINE PRICING, PRICE DISPERSION, AND TICKET CHARACTERISTICS ON AND OFF THE INTERNET

Overview

This study uses a unique, individual transaction level data set to contribute to our understanding of the impact of internet purchases on airline ticket prices. The analysis considers both the direct effect of internet purchase on prices paid and the effect of increased route shares of internet purchases on the level and dispersion of prices more generally. The analysis also extends previous work by Borenstein (1989) and Borenstein and Rose (1994) regarding how market concentration and hubbing affects the level and dispersion of airline prices. Our data enables us to investigate these issues while controlling for ticket characteristics and restrictions, time of purchase, and estimated load factors.

The sale of products on the internet has dramatically risen over the past decade. Internet sales compete directly with traditional outlets for the sale of a wide range of products. Consumers now buy numerous products and services online, including books and CDs, electronic products through shop-bots, and travel and entertainment. Internet penetration is particularly substantial in airlines where in particular city-pairs the internet share exceeds fifty percent of transactions. The internet theoretically reduces search costs and enables buyers to identify low price sellers, enhancing competition and reducing prices and price dispersion.

As noted by many, however, the empirical results supporting these theoretical predictions have not generally been strong or consistent, although data limitations have played an important role in limiting the analysis of different markets. Early studies found mixed evidence regarding prices and price dispersion.[2] Morton, Zettlemeyer, and Silva-Rissio (2002) use micro level transaction data and they find that consumers who searched for automobiles online paid about two percent less than consumers who did not. Brown and Goolsbee (2002) investigate the effects of internet search on the prices of term life insurance. They compare prices in geographic areas and for demographic groups where there are high levels of internet search for life insurance with areas and groups where internet search is low. They find that the internet has lead to a substantial reduction in both the level and dispersion of term life insurance prices. Baye and Morgan (2001) use posted prices and investigate the role of shopbots and other middlemen, finding high levels of price dispersion on the internet. Similarly, Scholten and Smith (2002) examine cross-market variation and find greater price dispersion for internet purchases for a wide variety of goods as compared to prices drawn from heterogeneous geographical settings.[3]

These studies provide important information regarding varied aspects of internet pricing, but due to data limitations they do not provide a comprehensive analysis of the impact of the internet on internet prices, overall price levels, and price dispersion for an entire industry. Instead these analyses are each drawn from different industries, and

[2]Bailey focused on books, CDs and software (Bailey (1998)), finding higher online prices and equal price dispersion across these channels. Brynolfson and Smith (2000) focus just on books and CDs and find lower prices but find that internet price dispersion is quite high.
[3] Chen (2006) finds that average prices on the internet for airline tickets differ by about 3 percent among the online travel agents, consistent with the findings of Clemons et al. (2002).

each investigates a particular aspect of how the internet has affected prices. Further, the data used often consist of posted prices rather than transaction prices.[4] These analyses also at times compare "national" internet prices with offline prices drawn from potentially heterogeneous local markets.

This research provides to our knowledge the most comprehensive analysis of the impact of internet purchases on pricing for a single industry, using a unique data set from the airline industry.[5] The central goal of the paper is to investigate the effects of internet sales on prices paid for airline tickets.

The data set consists of actual individual transactions purchased through a large Computer Reservation System (CRS).[6] These actual transactions data offer substantial advantages over data collected from posted prices because they include actual purchases and reflect substantial heterogeneities in units purchased for tickets with different characteristics and prices. The data include whether a ticket was purchased online or offline and numerous observable ticket characteristics and restrictions, such as refundability, advance purchase requirements and travel restrictions. The analysis also controls for estimated load factors at the time of purchase, network peak times, market

[4] Note that much of the existing literature on internet pricing relies on posted prices as compared to transaction data (Pan et al. (2003)).
[5] Note that Verlinda and Lane (2004) has analyzed the effects of the internet on airline pricing. Their analyses match data from DB1B with data measuring the average share of people using the internet to purchase airline tickets in a given geographical area. These analyses are limited because they do not use observed internet transactions. More important, their analyses do not consider route-level variation in internet share from a given origin, and they also do not have available other ticket characteristics of the type considered here. The analysis below shows that internet savings will be overestimated if one does not control for ticket characteristics.
[6] These data were matched with ticket characteristics using a procedure described in more detail below.

structure, and route characteristics.[7] The results show that these factors account for the large majority of observed variation in ticket prices; a regression of individual ticket prices on these characteristics, route, and carrier dummies yields an R^2 of more than .8.

Our analysis measures the direct impact of internet purchase for those customers who buy on the internet, controlling for ticket and market characteristics. We also investigate how increased shares of internet purchases influence both average prices online and offline and examines the impact of internet purchases on the dispersion of prices for online fares, offline fares, and fares overall. Finally, the analysis provides a more detailed investigation of the relationship between market structure and price levels and dispersion, controlling for ticket and flight characteristics.

The results show that when controlling for ticket characteristics, tickets purchased on the internet cost about 13 percent less than tickets purchased offline. The internet also enables consumers to find the lowest-priced package of characteristics. If one does not control for ticket characteristics, estimated internet savings rise to more than 40 percent.[8] The results also show that high levels of internet purchases drive down both offline and online prices by about ten percent, and these effects are actually larger in the offline market. Hence internet shopping in this industry generates a general market benefit similar to that underlying the results of Brown and Goolsbee. The results

---

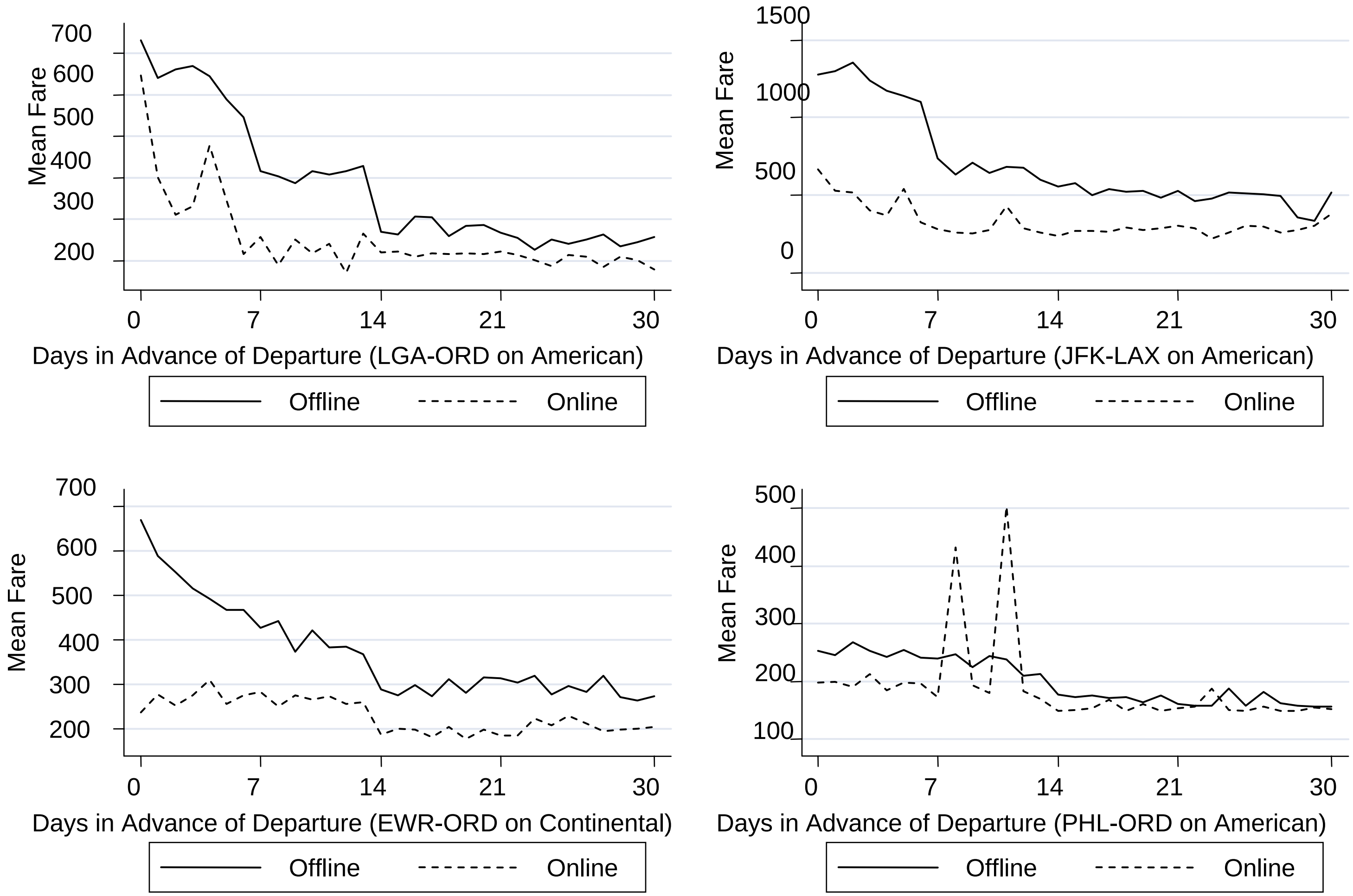
[7] The data were provided by one of the major computer reservation systems (CRS) and include individual transaction data for the fourth quarter of 2004. The data indicate city-pair, carrier, flight number, dates of purchase, departure, and return, booking class, and whether the ticket was purchased online or offline. These data were then matched with contemporaneous data from another CRS that included full details on individual ticket characteristics such as refundability, booking class, advance purchase requirements, travel and stay restrictions, and Saturday night stay-over.

[8] Note that this result is driven in part by the adverse selection problem where customers who want lower fares and are willing to accept various restrictions are more prone to shop on the internet.

also show that greater internet penetration reduces price dispersion in the market as a whole by driving offline prices down toward internet prices and by reducing dispersion within the online and offline segments.

Figure 1 presents some illustrative data. The figure shows mean fares measured for various days prior to departure for online versus offline sales for the four largest city-pair/carrier combinations in our data. These data indicate that internet purchasers pay substantially lower fares than offline purchasers, regardless of the number of days in advance when purchase occurs.

Table 1 provides a further look at these data, investigating the general relationship between online and offline fares for tickets containing various restrictions. The table shows average online and offline fares for the largest carrier on our eight largest city-pairs for tickets with various characteristics. The characteristics considered include class of service, refundability, advance purchase requirements, Saturday stay-over, and other ticketing restrictions. The results show that internet fares are consistently significantly lower in these categories than offline fares. The analysis below provides an in-depth investigation of these differences.



Mean Fares Offline and Online for Purchases in Last 30 Days Before Departure

Figure 1
Comparison Between Online and Offline Daily Average Fares

Table 1
Comparison of Online and Offline Mean Fares Conditional on Ticket Characteristics

	Delta		**US Airways**		**Continental**		**American Airlines**			
	Atlanta-La Guardia		**Boston-Dulles**		**Los Angeles-Newark**		**Boston-Chicago**		**Los Angeles-JFK**	
	Offline	Online	Offline	Online	Offline	Online	Offline	Online	Offline	Online
Ticket Characteristics:										
3-day advance	351.68	240.73	691.42	682.00	492.13	370.00	396.48	308.20	1549.47	553.79
7-day advance	264.77	233.47	384.08	242.27	307.24	283.11	269.34	198.23	336.70	264.41
14-day advance	191.28	185.20	300.20	239.93	293.76	292.18	292.06	202.78	361.54	283.05
Refundable	1365.12	426.00	435.44	394.46	1869.45	NA	1029.91	NA	1866.56	1523.33
Non-refundable	276.87	216.11	275.28	211.91	490.72	331.13	384.24	230.20	444.39	302.50
Saturday stay-over	239.18	202.29	294.63	225.13	478.96	299.67	268.07	205.13	643.37	283.46
Business/Full coach class	1105.44	NA	680.67	652.16	782.68	386.33	916.78	NA	1448.76	NA
First Class	1360.04	437.99	736.10	304.50	2128.44	NA	737.69	NA	2340.96	NA
Minimum stay	181.49	174.99	235.12	207.73	360.93	309.94	276.98	204.30	349.72	274.11
Maximum stay	181.48	175.03	274.65	211.61	440.84	412.30	281.11	204.56	421.27	347.28

Table 1, Continued

	American Airlines		Continental		American Airlines			
	Chicago-DFW		Chicago-Newark		Chicago-Philadelphia		La Guardia-Chicago	
	Offline	Online	Offline	Online	Offline	Online	Offline	Online
Ticket Characteristics:								
3-day advance	587.15	1337.99	699.83	364.00	268.44	178.00	636.10	291.29
7-day advance	523.88	306.78	375.97	240.06	213.87	186.31	355.50	176.48
14-day advance	239.41	229.56	262.30	208.51	170.44	155.16	266.33	228.74
Refundable	1281.08	1223.99	1029.14	223.00	248.31	227.42	1182.55	NA
Non-refundable	561.74	308.04	383.18	235.11	199.76	158.70	407.30	219.59
Saturday stay-over	325.06	215.70	257.7	217.07	208.56	181.85	271.08	203.33
Business/Full coach class	849.48	1086.23	971.00	332.14	715.24	NA	1127.86	NA
First Class	1231.31	NA	1655.07	NA	620.55	628.00	1247.98	NA
Minimum stay	243.82	220.88	292.53	222.63	224.08	175.66	304.06	215.51
Maximum stay	235.50	210.61	287.83	230.67	248.67	172.10	308.96	221.24

The paper also contributes significantly to our understanding of the airline industry by providing a more complete analysis of how market structure affects the level and dispersion of airline prices. The analysis builds on the work of Borenstein (1989) and Borenstein and Rose (1994), who analyze the effects of market structure on price levels and price dispersion for airlines. Borenstein shows that hub dominance and high route shares lead to higher average prices. Borenstein and Rose show that price dispersion decreases as routes become more monopolistic. Due to data limitations neither study controls for important ticket characteristics or flight load factors, which have a significant effect on both price levels and price dispersion. Similar limitations apply to most existing studies of airline pricing. [9] Our analysis investigates the effect of market structure variables while controlling for internet purchases and ticket characteristics, purchase days in advance, estimated flight level load factors, and other peaking variables. Our results generally confirm those of Borenstein and Rose.

Dynamics of Airline Pricing

Airlines offer a wide variety of different fares for travel on the same flight and the same day. The available evidence indicates that airlines offer tickets for sale in a conceptual series of “bins” or “buckets,” where a bucket is defined by a series of ticket characteristics including class of travel, refundability, advance purchase requirements, and travel and stay restrictions such as minimum and maximum stays and/or Saturday

[9] Note that Stavins (2001) overcomes a part of this limitation by using some ticket characteristics in explaining the relationship between market concentration and prices in the airlines market. Stavins, however, uses certain posted prices and only a subset of ticket characteristics. Using these posted prices, Stavins shows that price discrimination decreases as market concentration rises while increases in the route share of a carrier allows it to price discriminate more among the consumers.

stay-over.[10] The received wisdom is that airlines limit the quantity of low price tickets by limiting the number of tickets in low price buckets. For example, certain combinations of characteristics may only be used during certain days of the week (e.g. TWF), and certain tickets may only be available for round trips. Certain fares may not be available on certain flights for a period of time, and then later become available. High priced tickets are sometimes sold far in advance of departure, and deeply discounted tickets in certain bins may be available on the day of departure.

Airlines can alter the prices passengers ultimately pay for tickets both by changing the price of tickets within a given bucket and by rationing the number of tickets in that bucket.[11] The general analysis of this issue is beyond the scope of this paper.

For the present analysis it is simply important to note that airlines price using these ticket characteristics, which implicitly place tickets in particular bins that feature different prices. The analysis below shows that variation in ticket prices is driven largely by variation in ticket characteristics in that a simple regression of price on ticket characteristics, carrier and route dummies explains roughly 80 percent of the variation in ticket prices.

Airline customers and travel agents search for airline tickets by attempting to find sets of characteristics the customer is willing to accept at the lowest possible price. The most important component of this search, in terms of its impact on the ultimate price, is to find an open "bucket" with acceptable characteristics that has a low price.

---

[10] See Smith (2001).
[11] See Smith (2001).

An empirically smaller effect is found by identifying low priced tickets within a given bucket. The analysis below separately identifies internet price reductions that occur due to finding lower priced buckets and from finding lower prices within a given bucket. It also identifies the externality of increased internet purchases in terms of driving down overall fares.

The search for low price tickets may take place either online where the customer directly investigates the fares offered by one or more online sites, or it may take place offline where the ultimate customer uses a travel agent.

Search Theory, Pricing and Internet

The analysis here takes the same approach to the theory of internet pricing as that found in Brown and Goolsbee (2002). We assume, as is implicit or explicit in much of the literature on the internet, that the internet lowers search costs. This assumption is consistent with the evidence presented below.

More formally, the analysis builds on the search model of Stahl (1989). Stahl assumes that a certain exogenous share of customers are fully informed regarding all prices available in the market, and that another group of customers must pay a search cost for each price quote received. Because customers search sequentially in the Nash equilibrium, stores choose prices from a price distribution rather than using a pure strategy. Searchers with positive search costs stop searching endogenously whenever the price they observe is at or below their endogenously determined reservation price.

Fully informed customers have no search costs and search exhaustively, buying from the lowest price seller.

This logic generates three well-known results noted by Brown and Goolsbee. First, in equilibrium firms draw prices from a distribution, generating price dispersion. Second, as the share of informed customer rises the price distribution shifts monotonically downward leading to decreases in average prices. Third, the degree of price dispersion is not monotonic with respect to the percentage of informed consumers. Instead, when there are no searchers there is a degenerate distribution at the monopoly price. Dispersion then rises as some consumers become informed. Dispersion eventually reaches a peak before descending back to zero when all consumers are fully informed and prices converge to the competitive level.

We will assume that online customers are better informed than offline customers, although ultimately we will investigate this assumption in the empirical analysis by determining if they pay lower prices holding ticket characteristics constant. Further, the analysis uses transactions prices rather than posted prices. The use of transactions prices means that we measure the distribution of actual transactions rather than the posted prices of sellers, some of which might net few if any sales.

It is also important to note that the data we use includes only fares available both online and offline. The data we used incorporated observations where transaction fares were matched to those found offline in an offline CRS. Further, the available information indicates that in general the same prices and fare combinations are available online and offline.[12] The analysis below investigates this issue by examining how the percentage of internet purchases affects the level of fares both on and off the internet.

Empirical Analysis

Data

This study uses a unique data set consisting of contemporaneous online and offline transaction data of airline tickets for the last quarter of 2004.[13] These data were provided by a leading Computer Reservation System (CRS) vendor and include all transactions for a large number of domestic routes handled by the CRS during that quarter. The CRS offers services to airlines, travel agents, and numerous online sites, so that the data include transactions for all three outlets, though we believe the share for airline sites is small. As noted above, the data from the CRS includes airline and flight number, origin and destination, fare, booking class, a fare code, and dates of purchase, departure and return. Overall, these data represent roughly thirty percent of domestic tickets sold. These data do not include refundability, advance purchase requirements, and travel and stay restrictions.

To obtain these variables, we electronically matched the data with a separate data set from a different CRS containing both fares offered and purchased for travel in particular city-pairs, by departure dates on particular airlines. These data included the

[12] Note that in the early days of Orbitz the prices listed there included pricing specials offered ‘directly’ by the airlines, falling outside the travel agent contracts. It may also be possible that some offered prices on particular airline sites that are lower than fares offered elsewhere. However, this paper does not include web special fares by virtue of the construction of the data set. See the Appendix.
[13] The data and construction of variables are discussed at length in Appendix A.

ticket characteristics not available from the first data set.[14] The data set from the second CRS was incomplete in that certain fares had been deleted from the archive, and so we were only able to match the fares imperfectly.[15] The criterion used was to keep transactions if we were able to match the fares within 2 percent; for multiple matches within two percent we kept the closest.[16] The resulting data set contains individual ticket transactions that include ticket characteristics and restrictions, together with carrier, flight information, and dates of purchase, departure, and return.[17]

This procedure matches roughly 35 percent of the observations from the first data set. For both the online and offline transactions, our match rate is somewhat lower for the lowest priced tickets. Match rates for different city-pairs are illustrated in Figures 2-4. For example, Figure 2 shows the matches for Chicago to Newark; Panel A shows matches for all airlines, and Panel B for Continental, the market leader. Figures 3-4 show similar kernel densities for two other large city-pairs. The kernel densities show an under-representation of the very lowest fares for both all airlines and for the largest airlines on a route.

---

[14] We have been informed that fares offered on the various CRSs are normally the same, but that at times a fare will only be offered on some CRSs. This permits the use of departure dates to match the route, carrier, fares and fare classes in the first data set with the detailed ticket characteristics found in the second data set. The details are provided in the appendix.

[15] The data in the second archive are kept for unknown intervals of time. Individual fares are then deleted in an unknown pattern over time.

[16] The Appendix Table A3 reports results using a 5 percent matching criterion. Those results are qualitatively similar to the results reported below.

[17] Since the CRS de-regulation in 2004, the airlines are free to provide different fares to any distribution channel including the major CRSs, their own CRS and web-site and online travel agencies like Expedia. This necessitated the adoption of the matching rule(s) as discussed in the paper. Please refer to the appendix for discussion on the matching procedure.

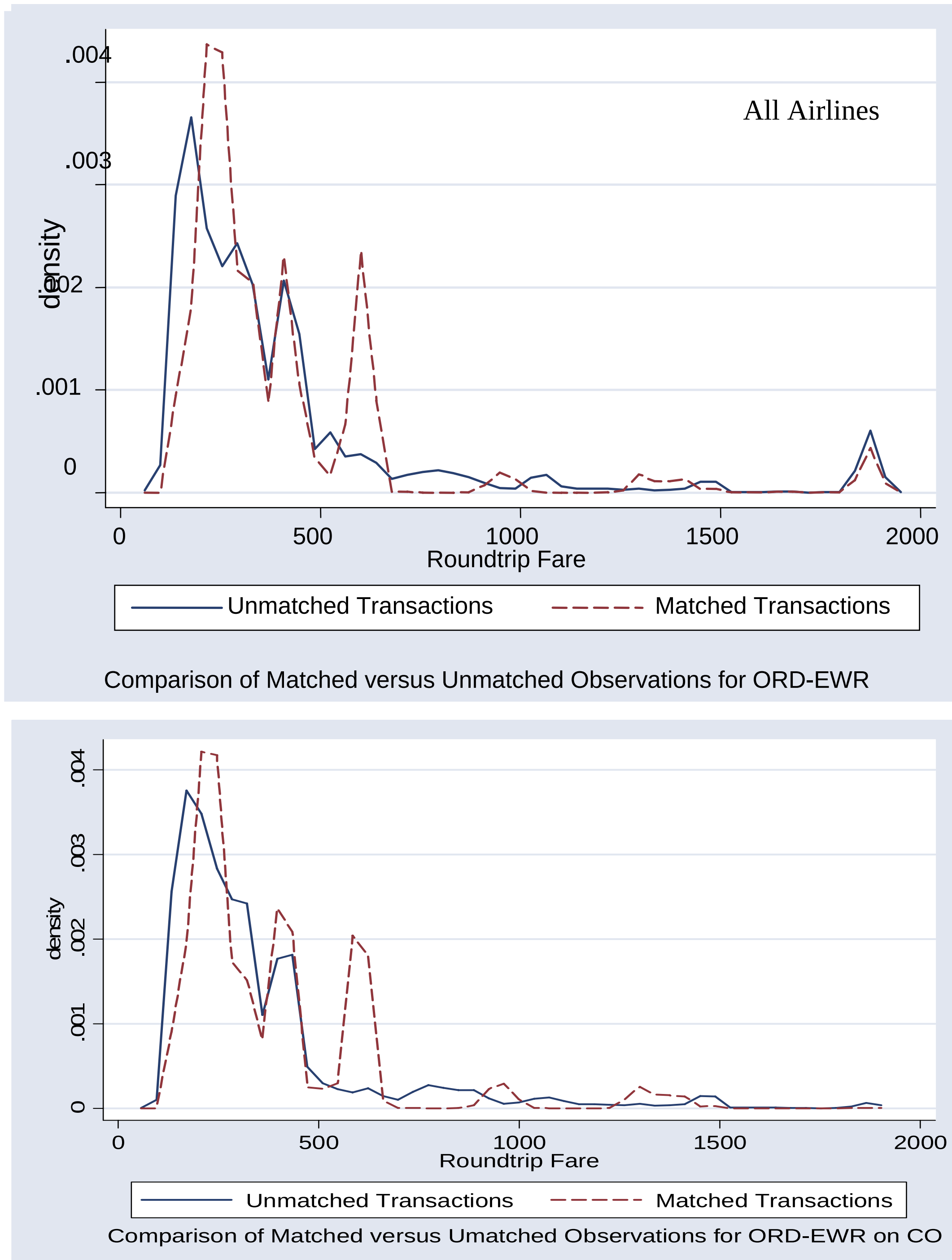
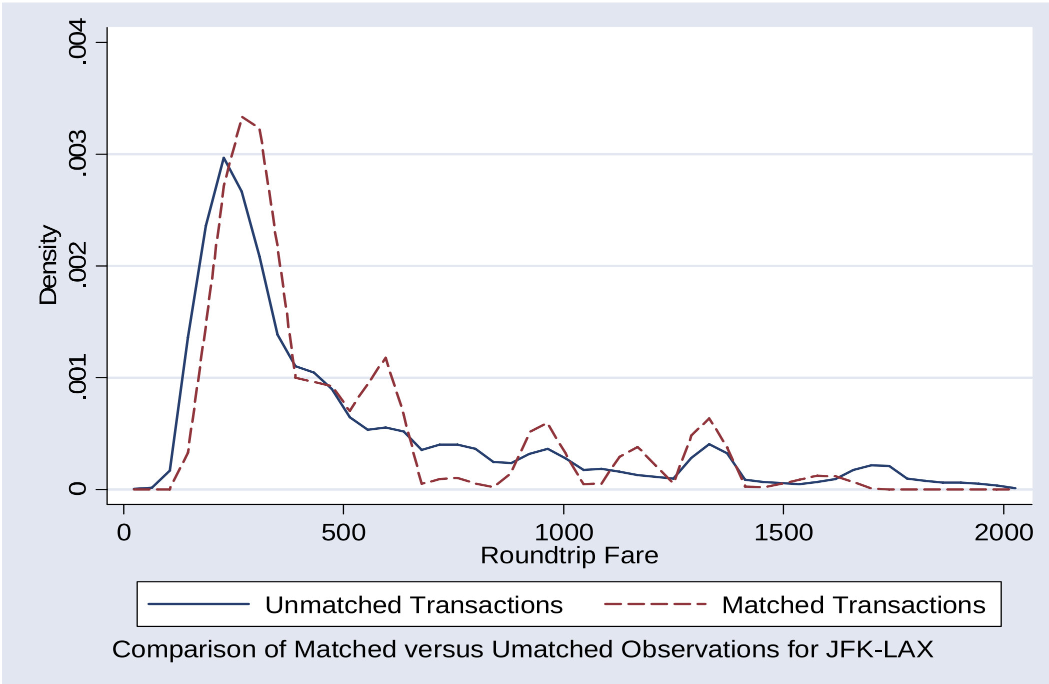


Figure 2
Comparing the Kernel Densities of Matched and Unmatched Transactions for flights between Chicago O' Hare – Newark Liberty



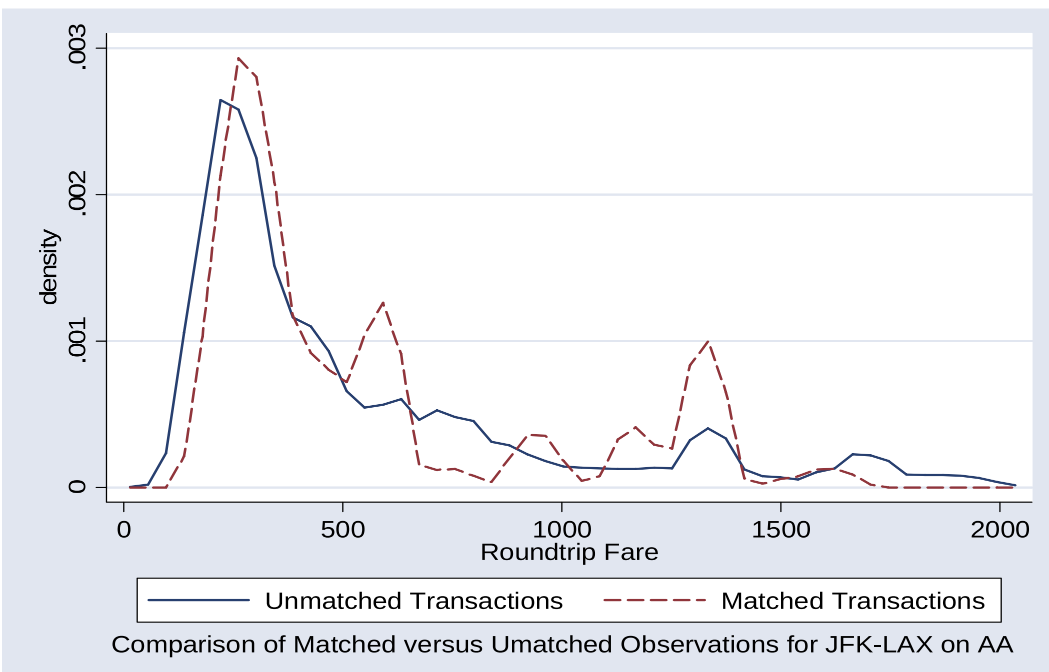
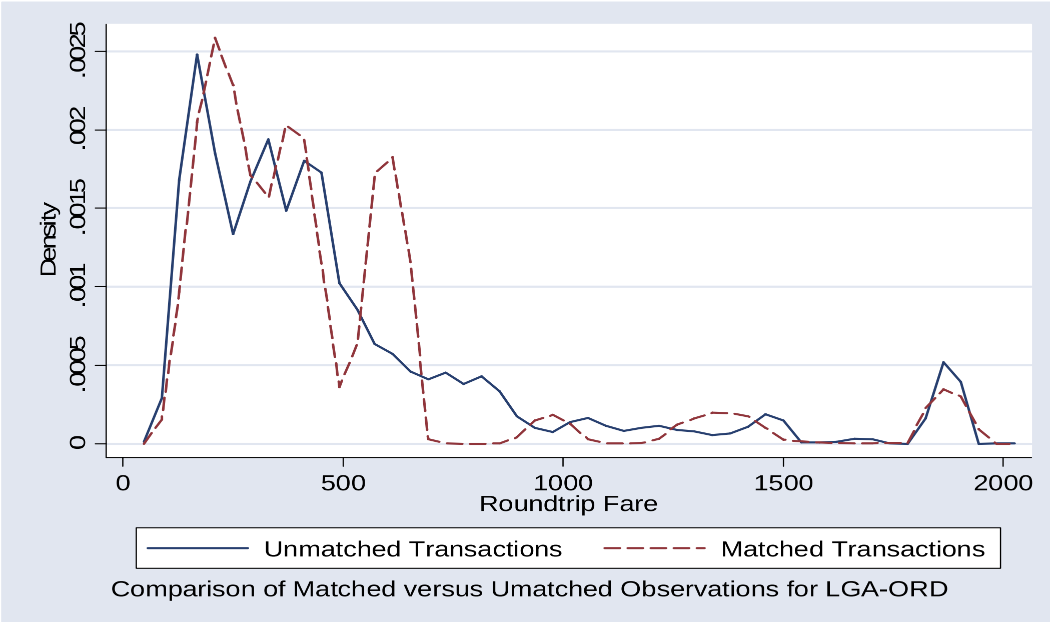


Figure 3
Comparing the Kernel Densities of Matched and Unmatched Transactions for flights between Kennedy, New York – Los Angeles, California



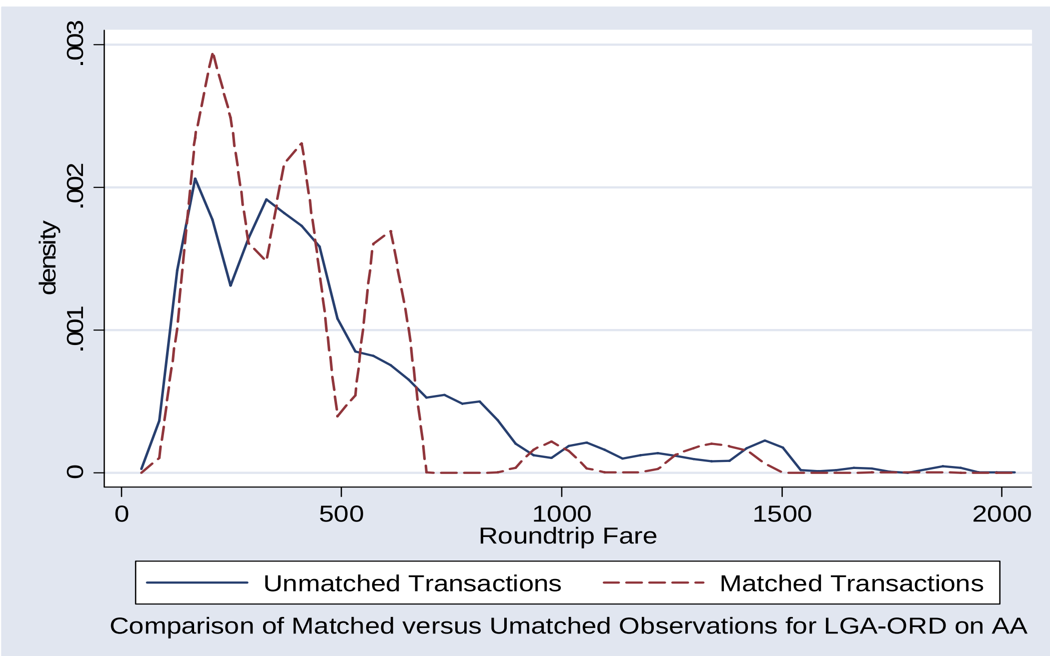
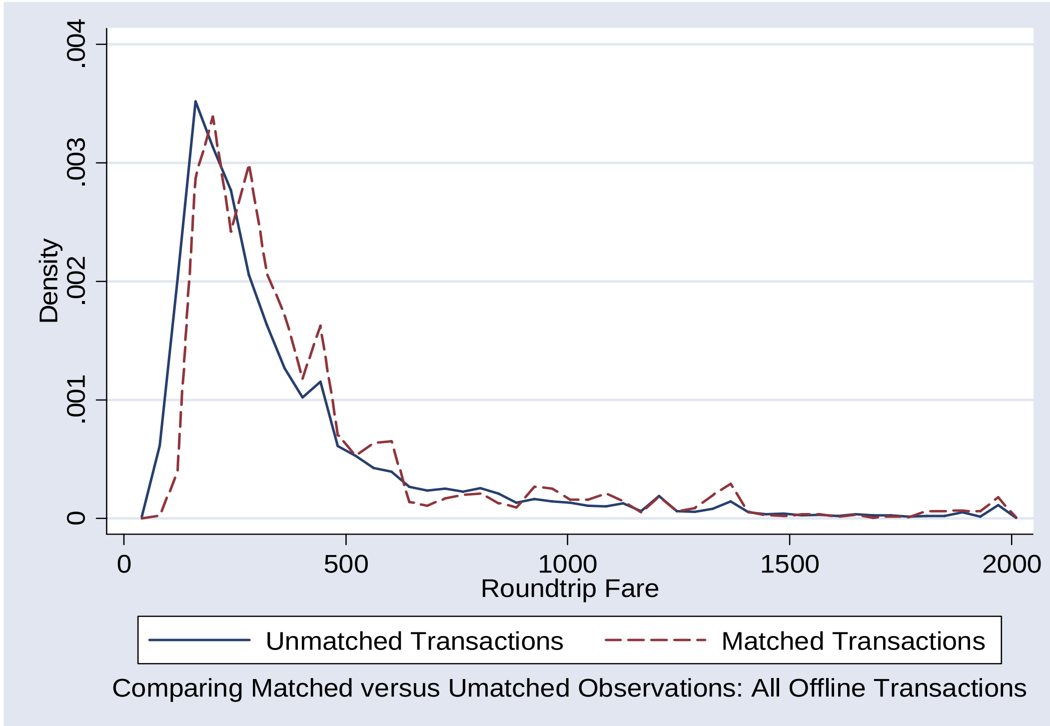


Figure 4
Comparing the Kernel Densities of Matched and Unmatched Transactions for flights between La Guardia, New York – Chicago O' Hare

Our analysis of online and offline fares, however, does not appear to be affected because we only consider matches for online and offline fares, and the under-representation of matches is comparable in both data sets. More specifically, Figure 5 compares the kernel densities for matched and unmatched transactions broken down by online versus offline transactions. Both online and offline transactions have fewer matches for very low fares, but there do not appear to be significant differences in the match rate for online versus offline fares.

The small difference in online versus offline matches is also illustrated by examining the match rate in the left hand tail of the distributions in Figure 5. This tail consists of price observations below $221, the price at which the matched versus unmatched kernel densities cross. Below $221, the match rate for offline tickets is about 22.6 percent while the match rate for online tickets is about 19.1 percent. These match rates are very comparable, and the differences point toward a small oversampling of offline fares. These effects point toward a comparative under-representation of low fares on the internet, although these effects will be minimized in the regressions because we control for ticket characteristics. We return to this issue below.

This study uses data for 150 U.S. domestic city-pairs (routes), including a mix of both business and tourist routes, and routes with varying groups of customers. A complete list of routes is contained in the Appendix, Table A2. We define a route as a city-pair regardless of direction.



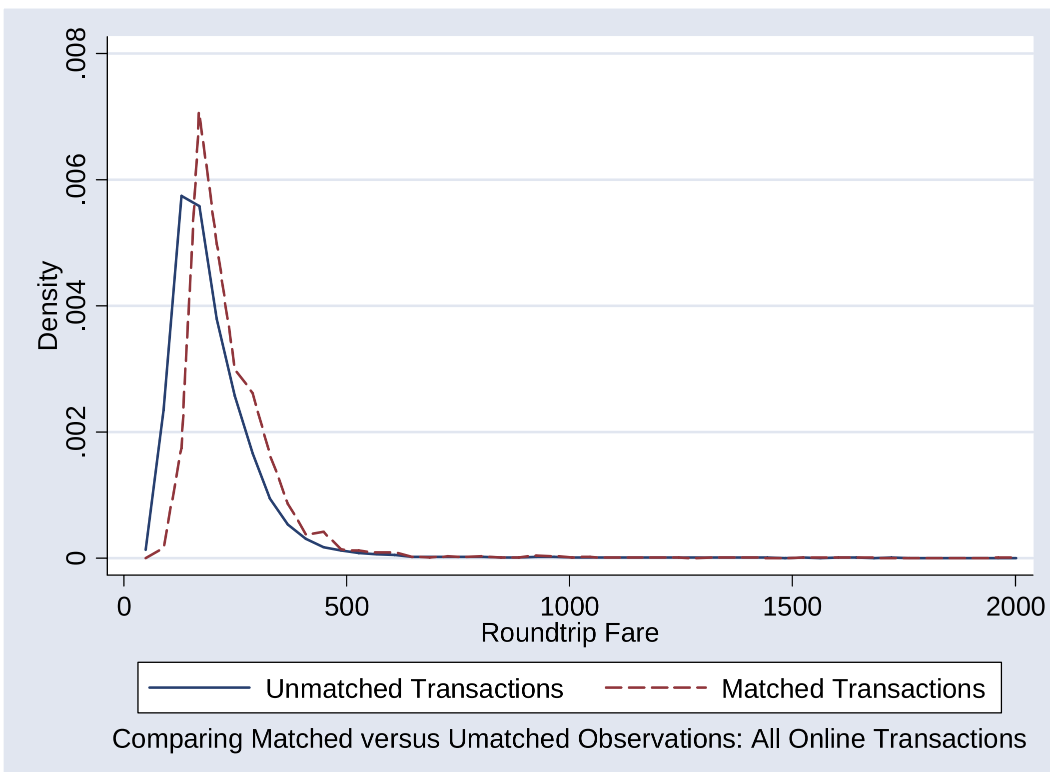


Figure 5
Comparing the Kernel Densities of Matched and Unmatched Transactions for Overall Online and Offline Transactions

Following Borenstein (1989) and Borenstein and Rose (1994), we include itineraries with at most one stop-over in either direction. The prices used are for roundtrip fares, doubling the fares for one-way tickets to obtain comparability. We exclude itineraries with open-jaws and circular trip tickets. This study includes tickets for flights operated by American Airlines, Continental, Delta, Northwest, US Airways, United Airlines, Frontier, Air Tran, Spirit, Alaska, American Mid-west, Sun Country, Hawaiian Airlines and American Trans Air.[18]

Observations consist of individual tickets, their fares and characteristics, and other data described above and more fully in the Appendix. The data also includes control variables for carrier and route effects, route market shares, HHI, hubs, and other standard variables measuring tourism, income, and population. We also include variables indicating the presence of discount carriers on routes, and a separate variable for Southwest. The complete set of variables is discussed later in this section.

Estimation Methodology

The basic model to be estimated regards the relationship between internet purchase and prices, which is estimated using the following specification:

$$\begin{aligned}\ln(Price_{ijtk}) = \beta_0 + \beta_1(R_{ijtk}) + \beta_2(Mktstructure_j) + \beta_3(Mktstructure_{jk}) + \beta_4(Online_{ijtk}) \\ + \beta_5(Loadfactor_{ijtk}) + \beta_6(C_{ijtk}) + \beta_7(DEP_{ijtk}) + \beta_8(RET_{ijtk}) + \varepsilon_{ijtk} \quad (1)\end{aligned}$$

where i refers to the ith ticket, j represents route j, t represents day t (departure date) while k stands for the kth carrier.

---

[18] We can identify routes served by Southwest, but we do not have data regarding Southwest's ticket characteristics because they are not included in one of the data bases.

R_{ijtk} represents the vector of the ticket restrictions (or characteristics) associated with the i^{th} ticket. The set of characteristics includes sets of dummy variables including refundability, advance purchase requirements, Saturday night stay-over, direct routes, round trip travel, first class, business class, travel restrictions (for example, if the ticket is valid only if you travel on a MTF), and stay restrictions (minimum or maximum stay requirement). These variables also include the number of days in advance the ticket was purchased. These variables also include peak times of day, such as flights departing on weekdays between 7:00 and 10:00 a.m. or between 3:00 and 7:00 p.m.[19] These are periods of peak travel, generally with high load factors and there may be network cost/capacity effects resulting in higher prices during these periods. For similar reasons, certain days of the week are busier than others so we include a full set of dummy variables representing each day of the week, with separate variables included for the departure day of the week and the return day of the week. Sunday is the omitted day.

The variables $Mktstructure_j$ represent a vector of route specific variables that have been widely used to study airline pricing (see, e.g., Borenstein). The Herfindahl index is measured using passenger shares in the fourth quarter of 2004 from DB1B. We also use a low cost route dummy indicating the presence of a low cost carrier and a separate Southwest dummy indicating Southwest's presence. We also include the logs of statutory distance, average population of the cities in the city-pair, average per capita income, and the temperature differential between the origin and destination, with the latter potentially measuring tourist effects (see Brueckner and Spiller (1991, 1994)). The

[19] We also used the time window of 8-10am as a peak time window. The results were qualitatively unaffected.

second market structure variable $Mktstructure_{jk}$ consists of the market share of carrier k on route j and a hub dummy if the carrier has a hub at the origin or destination. $Online_{ijtk}$ is an indicator variable corresponding to online purchases; C_{ijtk} represents carrier-specific fixed effects.

We employ both OLS and Instrumental variable (IV) regression techniques to estimate equation (1). The IV regression is required to address potential endogeneity issues of market share and the Herfindahl index. In the IV approach, we use the same instruments as Borenstein (1989) and Borenstein and Rose (1994). The carrier's market share is endogenous because it is influenced by prices and is instrumented using the carrier's enplanement share at the two endpoints on the route.[20] To the extent that market share is endogenous, the route Herfindahl index is also endogenous since the square of market share is one component of the route Herfindahl. The instrument for the Herfindahl index is the square of the fitted value for market share (from its first stage regression) plus the 'rescaled' sum of the squares of all other carriers' shares on the route.[21]

To investigate the effects of increased internet usage on fares generally we include an internet share variable, representing the ratio of online transactions to total transactions on a route. This measure is conceptually similar to that used by Brown and Goolsbee (2002) except that they use an estimate of people who have looked at prices on the internet for a given group, while our data permit direct measurement of the share of

[20] The Herfindahl index is subject to endogeneity issues because it is composed of market shares.
[21] See Borenstein (1989) and Borenstein and Rose (1994) for further discussion.

transactions completed on the internet.[22] We also include an interaction variable to measure whether the savings from internet purchase vary across markets with the level of internet usage.

Data Overview

Table 2 presents the descriptive statistics of the variables used. The final data set consists of 523,618 observations from 150 major routes in the US.[23] Our data contains two measures of internet usage. The first variable, online, is a dummy variable that takes a value of 1 if the transaction was completed over the internet. The second variable, internet share, is the ratio of all tickets purchased on the internet to all transactions on a given route. The share of internet purchases varies significantly across routes, varying from a meager 2 percent to as high as 58 percent. As expected, a higher share of purchases on the internet is found on leisure routes such as those to and from Orlando and Las Vegas.

[22] Recall from above that in our matched data all fares are available both online and offline.
[23] Please refer to Appendix A for a complete list of the routes included.

Table 2
Descriptive Statistics

Variable Description	**Mean**	**Standard Deviation**	**Minimum**	**Maximum**
Round-trip fare	474.840	466.065	61.99	4647.99
Non-refundable	0.756	0.429	0.000	1.000
Advance purchase requirement	5.428	6.149	0.000	30.000
Days prior to departure ticket purchased	15.514	20.050	0.000	202.000
Saturday stay-over	0.173	0.378	0.000	1.000
Travel restriction	0.398	0.489	0.000	1.000
Minimum stay requirement	0.218	0.413	0.000	1.000
Maximum stay requirement	0.190	0.392	0.000	1.000
First class	0.080	0.272	0.000	1.000
Business or Full coach fare class	0.122	0.327	0.000	1.000
Online	0.123	0.329	0.000	1.000
Internet share	0.184	0.096	0.024	0.587
Direct flight	0.989	0.103	0.000	1.000
Roundtrip	0.728	0.444	0.000	1.000
Load factor at time of purchase	0.113	0.084	0.003	0.880
Absolute temperature difference	15.217	10.893	0.001	46.000
HUB	0.742	0.437	0.000	1.000
Slot constrained airport	0.263	0.440	0.000	1.000
Departure at peak time	0.296	0.456	0.000	1.000
Return at peak time	0.215	0.410	0.000	1.000
Low cost carrier on route	0.473	0.499	0.000	1.000
Southwest Airlines	0.061	0.240	0.000	1.000
Distance	957.551	632.932	185.000	2704.000
Average population*	1986562	1593612	233014.6	5974809
Average per capita income*	36530.760	3321.425	23808.000	45046.490
Market share	0.546	0.254	0.000	1.000
HHI	0.538	0.191	0.189	1.000

*In thousands
Source: Ticket characteristic and fare data comes from one of the major CRS.
Market level data including market share and HHI is calculated using Department of Transportation's T-100 database.
List of hub airports is gathered from Air Traveler's (http://www.faqs.org/faqs/travel/air/handbook/part2/section-13.html)
For data descriptions please refer to the Appendix.

The average load factor calculated at the time of purchase has a mean of 0.11. It needs to be emphasized that the average load factor is defined as the average load factor of the individual segments for an itinerary at the time a ticket was purchased. Further, we calculate the load factor using both matched and unmatched observations in our transactions data, and for through tickets where the given route only consists of a portion of the trip. We also include open-jaws and circular trips in calculating load factors. Since data for all transactions on a specific flight are not available, our measured load factors only account for about 30 percent of transactions, corresponding to the share of transactions contained in our transactions data base. This means our load factors will be less than one. First class tickets on average comprise 8 percent of our data while business class and/or full coach class tickets account for 12 percent of our final data.[24]

Table 3 presents a simple hedonic regression of individual ticket prices on ticket restrictions and characteristics. The regressors include the ticket characteristics described above, the scarcity measures described, a hub variable, and route fixed effects. The results show that these variables explain about 80 percent of the variation in ticket prices.

The regression coefficients generally have the expected signs, and the results comport with conventional wisdom regarding airline pricing. The results show that the ticket characteristics with the largest effects on ticket prices are refundability, class of travel, stay restrictions, and advance purchase requirements. Refundable tickets on

[24] These tickets do not include first class itineraries. The transaction data allows us to differentiate between first class and business or full *coach* class tickets. See Appendix for a detailed discussion on the same.

average cost about 44 percent more than non-refundable ones. Advance purchase requirements generally reduce ticket prices, sometimes by more than 50 percent. These large advance purchase requirement effects occur even though we separately control for the days in advance a ticket was bought. It is worth noting, however, that the pricing for advance purchase requirements is not monotonic.

Table 3
Regression of Natural Logarithms of Fares on Ticket Characteristics

	Log(Fare)
Advance Purchase Requirement:	
One day advance	-0.2497
	(68.16)**
Three day advance	-0.0441
	(26.72)**
Five day advance	-0.5551
	(39.17)**
Seven day advance	-0.1706
	(120.88)**
Ten day advance	-0.2079
	(67.18)**
Fourteen day advance	-0.2607
	(162.62)**
Twenty-one day advance	-0.3111
	(94.34)**
Thirty day advance	-0.1282
	(16.30)**
Other Characteristics:	
Non-refundable ticket	-0.4389
	(268.33)**
Days prior to departure ticket purchased	0.0003
	(9.62)**
Saturday stay-over	-0.1377
	(73.74)**
Travel restriction requirement	-0.2803

Table 3, Continued

	Log(Fare)
	(268.44)**
Minimum stay requirement	-0.0067
	(4.45)**
Maximum stay requirement	-0.0467
	(30.74)**
First class	0.7057
	(364.18)**
Business class	0.3205
	(194.56)**
Direct flight	0.0743
	(18.01)**
Roundtrip ticket	-0.1101
	(50.07)**
Flight load factor at purchase	0.2634
	(40.06)**
HUB	0.025
	(9.37)**
Peak Time of Day:	
Departure at peak time	0.0087
	(9.03)**
Return at peak time	0.0188
	(16.54)**
Departure Day of the Week (Sunday omitted):	
Monday	0.0051
	(3.16)**
Tuesday	0.0028
	(1.69)
Wednesday	0.0062
	(3.52)**
Thursday	0.0114
	(6.18)**
Friday	0.0159
	(8.09)**
Saturday	-0.0735
	(34.46)**
Return Day of the Week (Sunday omitted):	
Monday	-0.035
	(16.36)**
Tuesday	-0.024
	(10.67)**
Wednesday	-0.0293
	(12.94)**
Thursday	-0.0205
	(9.03)**
Friday	-0.0304
	(13.54)**

Table 3, Continued

	Log(Fare)
Saturday	-0.0716
	(26.34)**
Carrier Fixed Effects (American Airlines omitted):	
Continental	-0.0029
	(1.08)
Delta	-0.0072
	(3.11)**
Northwest	-0.0195
	(5.61)**
United Airways	0.0937
	(61.40)**
US Air	-0.0301
	(10.38)**
Frontier Airlines	-0.2562
	(63.52)**
Alaska	-0.0481
	(5.11)**
Hawaiian Airlines	0.0075
	(0.49)
America West	-0.231
	(56.98)**
American Trans Air	-0.2323
	(16.12)**
Midwest Express Airlines	-0.0487
	(3.23)**
Air Tran	-0.3601
	(33.32)**
Spirit	-0.2552
	(42.60)**
Sun County	-0.4961
	(58.01)**
Route effects	Yes
Constant	6.102
	(846.86)**
Observations	523618
R-squared	0.790

Longer periods of advance purchase do not necessarily correspond to lower prices. First class tickets cost about 70 percent more than ordinary tickets, and business class tickets cost about 32 percent more. The results also show that travel restrictions have a major effect on ticket prices, reducing ticket prices by about 28 percent. Minimum and maximum stay requirements are also associated with slight reductions in ticket prices. Direct trips involving no change of planes cost roughly 7 percent more than itineraries involving a plane change while roundtrip tickets are 11 percent less expensive. An itinerary involving a Saturday night stay-over on average costs roughly 14 percent less than an itinerary with similar characteristics that does not.

The peaking variables are all highly statistically significant and of the right sign, but their economic magnitudes are small. Leaving at a peak time only increases average fares by less than one percent. Returns are most expensive on Sunday (not a peak day), and again the effects are small. Itineraries that involve a departure and/or return during peak times of day increase fares by .8 and 1.9 percent respectively.[25] The results also show that a 10 percent increase in the average load factor at the time of ticket purchase increases fares on average by 3.0 percent. The results indicate that an increase in load factors at the time of purchase by one standard deviation increases fares by about 2.1 percent. Overall, these peaking variable coefficients are economically small.

Empirical Results

Direct Effect of Internet on Airline Prices

Turning to the effects of the internet, Table 4 presents the results testing our central hypothesis that internet purchases lead to lower prices. These regressions also

---

[25] These variables reflect peak travel times between 7:00-10:00 a.m. and 3:00-7:00 p.m., when there are generally high load factors on routes and throughout airline networks.

introduce market structure variables for additional control. These variables and other route-specific variables could not be included in Table 3 because of the route fixed effects. Table 4 present OLS and instrumental variable (IV) regressions of individual ticket prices on ticket characteristics, market structure variables and carrier fixed effects.[26] These regressions test the hypothesis that online consumers pay lower prices on average for tickets, controlling for ticket characteristics, route, and carrier.

The coefficients of the remaining control variables are similar to those in Table 3. The coefficients on market structure variables are consistent with the literature. Our results show that an increase of 10 percent in the market share of a carrier increases prices by 2 percent on average. The coefficient on the Herfindahl Index is negative and statistically significant as found in several previous airline studies.[27] Fares on routes involving a carriers' hub at either or both endpoints average 13 percent more than fares on routes that do not involving a carrier's hub.[28]

The presence of an airport with restricted slots at either endpoint increases fares by about 14 percent.[29] On the other hand, the presence of a low cost carrier, other than Southwest, decreases average fares by roughly 10 percent, while Southwest's presence decreases average fares by 16 to 19 percent.

---

[26] IV regression involves instrumenting market share and HHI variables by their respective instruments.

[27] This result, although not in keeping with common expectation, is consistent with the literature for empirical studies examining the determinants of profitability where both market share and market concentration are included as explanatory variables. In airlines, see Borenstein (1989) and Stavins (2001). For a more general discussion on negative relationship between market concentration and prices see Ravenscraft (1983), Mueller (1986). Also see Appendix A.

[28] We included separate dummy variables to control for origin and destination hub airports for the operating carrier individually. Results show that the presence of a hub at the origin airport for the operating carrier increases prices by 11 percent while the presence of a hub at the destination airport increase prices by 9 percent on average.

[29] JFK (New York), La-Guardia (New York) and DCA (Washington D.C) are the three airports which still have restricted slots. Chicago O' Hare (ORD) was included in this list until 2002 (Lee and Prado (2005)).

Table 4
Internet Purchase, Internet Share and Potential Savings from Purchase in High-Low Internet Usage Markets

	(1) Log(Fare) (OLS)	(2) Log(Fare) (IV)	(3) Log(Fare) (OLS)	(4) Log(Fare) (IV)	(5) Log(Fare) (OLS)	(6) Log(Fare) (IV)
Market Structure Variables:						
Market share	0.1182	0.2868	0.1208	0.2924	0.1203	0.2919
	(32.95)**	(27.54)**	(33.79)**	(28.14)**	(33.64)**	(28.06)**
HHI	-0.0055	-0.2527	-0.0236	-0.258	-0.023	-0.2574
	(1.25)	(22.34)**	(5.36)**	(22.86)**	(5.24)**	(22.76)**
HUB	0.1365	0.1216	0.1313	0.1147	0.1313	0.1147
	(85.77)**	(61.10)**	(82.62)**	(57.66)**	(82.61)**	(57.66)**
Slot constrained airport	0.1392	0.1422	0.1386	0.1396	0.1385	0.1395
	(80.11)**	(78.01)**	(80.01)**	(76.81)**	(79.96)**	(76.79)**
Internet Variables:						
Online	-0.1326	-0.1285	-0.1287	-0.1243	-0.1418	-0.1315
	(79.09)**	(75.72)**	(76.89)**	(73.38)**	(38.65)**	(35.36)**
Internet share			-0.4546	-0.4806	-0.4648	-0.4861
			(55.50)**	(58.04)**	(54.20)**	(56.20)**
Internet share*Online					0.0584	0.032
					(4.03)**	(2.19)*
Advance Purchase Requirement: (No advance purchase required omitted)						
One day advance	-0.3881	-0.383	-0.3716	-0.3672	-0.3715	-0.3671
	(93.54)**	(91.72)**	(89.61)**	(87.96)**	(89.56)**	(87.93)**
Three day advance	0.0516	0.0555	0.0554	0.0591	0.0556	0.0592
	(27.39)**	(29.31)**	(29.51)**	(31.31)**	(29.59)**	(31.35)**
Five day advance	-0.707	-0.7075	-0.6668	-0.6668	-0.6673	-0.6671
	(42.10)**	(41.98)**	(39.79)**	(39.65)**	(39.82)**	(39.67)**
Seven day advance	-0.135	-0.1346	-0.1338	-0.1335	-0.1336	-0.1335
	(82.16)**	(81.56)**	(81.63)**	(81.18)**	(81.54)**	(81.12)**
Ten day advance	-0.247	-0.2433	-0.2456	-0.2428	-0.2453	-0.2426

Table 4, Continued

	(1) Log(Fare) (OLS)	(2) Log(Fare) (IV)	(3) Log(Fare) (OLS)	(4) Log(Fare) (IV)	(5) Log(Fare) (OLS)	(6) Log(Fare) (IV)
	(67.98)**	(66.65)**	(67.79)**	(66.72)**	(67.71)**	(66.68)**
Fourteen day advance	-0.1948	-0.1913	-0.1979	-0.1952	-0.1978	-0.1951
	(105.28)**	(102.80)**	(107.24)**	(105.23)**	(107.11)**	(105.16)**
Twenty-one day advance	-0.2465	-0.2468	-0.2486	-0.2493	-0.2481	-0.2491
	(64.20)**	(64.02)**	(64.91)**	(64.88)**	(64.78)**	(64.79)**
Thirty day advance	0.0585	0.0567	0.057	0.0557	0.0577	0.056
	(6.35)**	(6.14)**	(6.21)**	(6.04)**	(6.28)**	(6.08)**
<u>Other Ticket Characteristics:</u>						
Non-refundable ticket	-0.3512	-0.356	-0.35	-0.354	-0.3499	-0.354
	(202.04)**	(203.22)**	(201.91)**	(202.74)**	(201.84)**	(202.69)**
Days prior to departure ticket purchased	-0.0002	-0.0003	-0.0001	-0.0001	-0.0001	-0.0001
	(7.23)**	(8.51)**	(3.23)**	(4.26)**	(3.54)**	(4.42)**
Saturday stay-over	-0.1296	-0.1298	-0.1269	-0.127	-0.1272	-0.1271
	(57.49)**	(57.37)**	(56.44)**	(56.32)**	(56.54)**	(56.36)**
Travel restriction requirement	-0.2839	-0.2853	-0.2801	-0.2807	-0.2801	-0.2807
	(237.53)**	(236.50)**	(234.62)**	(233.11)**	(234.60)**	(233.10)**
Minimum stay requirement	0.0099	0.0137	0.008	0.0113	0.008	0.0114
	(5.91)**	(8.14)**	(4.79)**	(6.77)**	(4.82)**	(6.78)**
Maximum stay requirement	-0.0121	-0.0088	-0.0098	-0.0066	-0.0097	-0.0066
	(7.07)**	(5.12)**	(5.72)**	(3.87)**	(5.69)**	(3.85)**
First class	0.7862	0.7831	0.7783	0.7746	0.7781	0.7745
	(352.69)**	(348.75)**	(349.47)**	(345.24)**	(349.27)**	(345.17)**
Business class	0.4256	0.4349	0.4326	0.4407	0.4326	0.4407
	(240.27)**	(240.46)**	(244.30)**	(243.58)**	(244.29)**	(243.55)**
<u>Remaining Ticket Characteristics:</u>						
Direct flight	0.0301	-0.0013	0.0247	-0.0066	0.025	-0.0064
	(6.14)**	(0.24)	(5.04)**	(1.28)	(5.10)**	(1.25)
Roundtrip ticket	-0.1089	-0.1101	-0.112	-0.1137	-0.1117	-0.1135
	(41.38)**	(41.60)**	(42.64)**	(43.08)**	(42.53)**	(43.01)**

Table 4, Continued

	(1) Log(Fare) (OLS)	(2) Log(Fare) (IV)	(3) Log(Fare) (OLS)	(4) Log(Fare) (IV)	(5) Log(Fare) (OLS)	(6) Log(Fare) (IV)
Flight load factor at purchase	0.1221	0.1042	0.1063	0.0875	0.1049	0.0867
	(15.87)**	(13.40)**	(13.85)**	(11.27)**	(13.66)**	(11.17)**
<u>Peak Time of Day:</u>						
Departure at peak time	0.0169	0.0181	0.017	0.0182	0.017	0.0182
	(14.66)**	(15.62)**	(14.76)**	(15.71)**	(14.77)**	(15.71)**
Return at peak time	0.0272	0.0282	0.0268	0.0278	0.0268	0.0278
	(19.96)**	(20.64)**	(19.75)**	(20.40)**	(19.74)**	(20.39)**
<u>Other Route Specific Characteristics:</u>						
Low cost carrier on route	-0.1019	-0.1025	-0.0831	-0.0829	-0.0829	-0.0828
	(85.93)**	(85.82)**	(67.61)**	(66.96)**	(67.38)**	(66.80)**
Southwest Airlines	-0.193	-0.1854	-0.17	-0.1619	-0.1704	-0.1622
	(77.74)**	(74.03)**	(67.72)**	(63.84)**	(67.83)**	(63.85)**
<u>Other Route Level Variables:</u>						
Distance (log)	0.3565	0.3519	0.3855	0.3816	0.3857	0.3818
	(352.07)**	(312.73)**	(339.22)**	(309.41)**	(338.98)**	(308.71)**
Absolute Temperature	-0.0022	-0.0024	-0.0014	-0.0015	-0.0014	-0.0015
Difference (Log)	(5.21)**	(5.67)**	(3.28)**	(3.50)**	(3.29)**	(3.50)**
Average population (Log)	-0.0289	-0.0304	-0.0378	-0.0389	-0.0378	-0.0389
	(35.60)**	(35.83)**	(45.90)**	(45.11)**	(45.86)**	(45.09)**
Average per capita	0.1375	0.1418	0.057	0.0575	0.0566	0.0572
Income (Log)	(20.68)**	(21.22)**	(8.41)**	(8.43)**	(8.34)**	(8.39)**
<u>Departure Day of Week (Sunday omitted):</u>						
Monday	0.0118	0.0121	0.012	0.0123	0.0119	0.0123
	(6.06)**	(6.22)**	(6.16)**	(6.31)**	(6.16)**	(6.31)**
Tuesday	0.0176	0.0183	0.0173	0.0178	0.0173	0.0178
	(8.74)**	(9.03)**	(8.59)**	(8.84)**	(8.58)**	(8.84)**
Wednesday	0.0223	0.0233	0.0223	0.0231	0.0222	0.0231

Table 4, Continued

	(1) Log(Fare) (OLS)	(2) Log(Fare) (IV)	(3) Log(Fare) (OLS)	(4) Log(Fare) (IV)	(5) Log(Fare) (OLS)	(6) Log(Fare) (IV)
	(10.54)**	(10.97)**	(10.56)**	(10.93)**	(10.53)**	(10.92)**
Thursday	0.0233	0.0248	0.0235	0.0248	0.0235	0.0248
	(10.58)**	(11.18)**	(10.69)**	(11.23)**	(10.67)**	(11.22)**
Friday	0.0288	0.0296	0.0284	0.0289	0.0285	0.029
	(12.19)**	(12.47)**	(12.04)**	(12.24)**	(12.08)**	(12.26)**
Saturday	-0.0547	-0.0546	-0.0545	-0.0545	-0.0545	-0.0545
	(21.36)**	(21.22)**	(21.33)**	(21.27)**	(21.34)**	(21.28)**
<u>Return Day of the Week (Sunday omitted):</u>						
Monday	-0.0349	-0.0349	-0.0345	-0.0344	-0.0348	-0.0346
	(13.61)**	(13.58)**	(13.48)**	(13.44)**	(13.61)**	(13.51)**
Tuesday	-0.0217	-0.0215	-0.0211	-0.0208	-0.0216	-0.0211
	(8.07)**	(7.95)**	(7.87)**	(7.73)**	(8.04)**	(7.82)**
Wednesday	-0.0312	-0.0311	-0.0299	-0.0296	-0.0304	-0.0298
	(11.52)**	(11.42)**	(11.07)**	(10.91)**	(11.22)**	(10.99)**
Thursday	-0.0218	-0.0224	-0.0206	-0.021	-0.021	-0.0212
	(8.02)**	(8.21)**	(7.59)**	(7.71)**	(7.74)**	(7.79)**
Friday	-0.0372	-0.0381	-0.0356	-0.0362	-0.036	-0.0364
	(13.82)**	(14.08)**	(13.25)**	(13.43)**	(13.39)**	(13.50)**
Saturday	-0.0756	-0.076	-0.0741	-0.0744	-0.0746	-0.0746
	(23.19)**	(23.24)**	(22.81)**	(22.82)**	(22.93)**	(22.88)**
<u>Carrier Fixed Effects (American Airlines omitted):</u>						
Continental	-0.0251	-0.0231	-0.0316	-0.0302	-0.0318	-0.0303
	(12.52)**	(11.51)**	(15.82)**	(15.06)**	(15.90)**	(15.10)**
Delta	-0.1275	-0.1365	-0.132	-0.1411	-0.132	-0.1411
	(72.24)**	(74.45)**	(74.88)**	(77.08)**	(74.90)**	(77.09)**
Northwest	0.0546	0.0439	0.0897	0.0807	0.0896	0.0806
	(23.37)**	(18.01)**	(37.19)**	(32.23)**	(37.16)**	(32.22)**
United Airways	0.1243	0.1167	0.1463	0.1412	0.1464	0.1413

Table 4, Continued

	(1) Log(Fare) (OLS)	(2) Log(Fare) (IV)	(3) Log(Fare) (OLS)	(4) Log(Fare) (IV)	(5) Log(Fare) (OLS)	(6) Log(Fare) (IV)
	(78.13)**	(72.07)**	(89.49)**	(85.37)**	(89.56)**	(85.39)**
US Air	-0.1166	-0.1324	-0.1005	-0.1166	-0.1007	-0.1167
	(52.77)**	(52.95)**	(45.22)**	(46.55)**	(45.29)**	(46.61)**
Frontier Airlines	0.0438	0.0538	0.0883	0.1023	0.087	0.1015
	(12.04)**	(14.34)**	(23.79)**	(26.65)**	(23.33)**	(26.30)**
Alaska	-0.1891	-0.2224	-0.1432	-0.1719	-0.1428	-0.1716
	(33.06)**	(37.68)**	(24.85)**	(29.02)**	(24.78)**	(28.96)**
Hawaiian Airlines	0.3237	0.2893	0.311	0.2799	0.3113	0.2801
	(19.80)**	(17.60)**	(19.08)**	(17.08)**	(19.10)**	(17.09)**
America West	-0.1256	-0.1218	-0.098	-0.089	-0.0979	-0.089
	(29.76)**	(27.53)**	(23.12)**	(20.05)**	(23.10)**	(20.04)**
American Trans Air	0.1351	0.1221	0.1419	0.1291	0.1423	0.1293
	(8.54)**	(7.69)**	(9.00)**	(8.15)**	(9.03)**	(8.17)**
Midwest Express Airlines	0.4837	0.6143	0.4693	0.5985	0.4686	0.598
	(31.18)**	(36.10)**	(30.34)**	(35.29)**	(30.29)**	(35.24)**
Air Tran	-0.315	-0.2592	-0.3519	-0.2988	-0.3522	-0.299
	(26.11)**	(20.84)**	(29.21)**	(24.08)**	(29.24)**	(24.09)**
Spirit	-0.1755	-0.1584	-0.127	-0.1061	-0.1294	-0.1074
	(28.60)**	(25.26)**	(20.56)**	(16.79)**	(20.84)**	(16.90)**
Sun County	-0.2016	-0.1352	-0.1612	-0.0925	-0.1632	-0.0936
	(23.18)**	(14.33)**	(18.53)**	(9.78)**	(18.73)**	(9.87)**
Constant	2.784	2.8811	3.6417	3.7644	3.646	3.7665
	(38.13)**	(39.26)**	(48.93)**	(50.26)**	(48.99)**	(50.28)**
Observations	523618	523618	523618	523618	523618	523618
R-squared	0.70	0.70	0.70	0.70	0.70	0.70

Note: Absolute value of t statistics in parentheses; * significant at 5%; ** significant at 1%; Source: Please refer to Table 1

The results also suggest that an increase in the temperature difference between the two endpoints on a route by 1 percent decreases fares by about 0.2 percent, which according to the literature captures a tourist effect. Distance between endpoints and the average per capita income at the endpoints increase average fares while increases in average population decrease average fares.[30]

Turning to the internet effects, the results show that after controlling for ticket characteristics, market structure, and carrier fixed effects online customers pay about 13 percent less than offline customers. Hence the results show very substantial effects of internet purchase on ticket prices. These results, which control for most sources of variation in ticket characteristics and market structure, are both economically and statistically significant.[31]

The regressions in Table 4 do not contain route fixed effects, but instead follow standard practice by including market structure variables. This standard procedure is required to allow estimation of the effect of market structure variables and other variables that do not vary at the route level, including the percentage of internet purchases.[32] Appendix A, however, contains a similar estimation with a full set of route effects. Such estimation requires omission of the market structure variables and other

[30] See appendix for further discussion of this variable.

[31] We have also re-run the regressions in Table 3 replicating the "bucket structure" described earlier in the paper; we included various combinations of restrictions to allow for variation in the pricing of different combinations of the restrictions described in Table 3. These regressions included the roughly 100 different restriction combinations that make up the large majority of tickets sold. The internet effects were roughly the same as in Table 3.

[32] Estimation without these effects follows well-established procedures for analysis in this industry, see for example, Borenstein (1989), Borenstein and Rose (1994), and Stavins (2001).

variables that only vary across markets, including these route fixed effects results in a reduced estimate of the direct effect of internet purchase of 11.4 percent.

The apparent reason for these observed price differences regards the superior search mechanism provided by the internet, combined with potential agency and/or communication problems with travel agents. As noted, fares consist of the pricing of tickets with a series of restrictions and the methodology ensures that all of the online fares were available both online and offline. The regressions control for characteristics accounting for roughly 80 percent of the observed fare variation. Yet even with such control, internet purchasers are able to find fares discounted about 13 percent compared to those purchased through travel agents. Individual customers searching on the internet are able to and do make tradeoffs in ticket characteristics that enable them to pay lower prices. Hence customers buying online find better deals within the various buckets.

The likely source of the higher offline fares consists either of communication or agency problems. While it appears that travel agents have access to the same fares, the process of communicating the large array of restriction combinations and the willingness of customers to accept particular restrictions is problematic. Regarding agency issues, travel agents do not face the same incentives as consumers who spend their own money on a ticket, and it may also be true that a greater percentage of offline tickets are reimbursed, resulting in offline customers using a different tradeoff between price and convenience.

Indirect Effect of Internet on Airline Prices

Columns (3) and (4) of Table 4 analyzes the effect of the share of internet purchases on the average level of fares. In addition to the variables used in columns (1) and (2) of Table 4, as discussed above, we include an internet share variable representing the share of internet purchases on a given route. The internet share variable is large and highly significant. The results show that a ten percent increase in the share of tickets purchased on the internet for a route, which is approximately one standard deviation of that variable, decreases average fares on the route by about five percent. These results capture a spillover of internet shopping onto prices that is consistent with the empirical findings of Brown and Goolsbee (2002).

Hence this specification shows that those who shop online receive direct benefits in terms of lower prices and also lower the prices of others who shop less.

This finding raises the issue of whether the returns to shopping are greater or smaller in markets where there are many shoppers. In Stahl's model, of course, shoppers obtain the competitive price because they know all prices. More generally, one can think of shoppers obtaining better prices than non-shoppers, but the precise returns would vary with the conditions of the market, the percentage of shoppers and the level of price dispersion.

Columns (5) and (6) of Table 4 present a partial analysis of this issue. The regressions in these columns investigate the relative magnitude of the savings of internet purchase as influenced by the share of total purchases made on the internet. More specifically, these regressions interact internet purchase with the online share variables.

The results indicate that there are small but significant differences in the level of internet savings when there are more internet purchases. As the share of tickets purchased on the internet rises by 10 percent (roughly one standard deviation), online savings fall by about 0.3 percent. This result suggests that as the share of internet purchases rises, the difference between the fares paid by the online and offline consumers falls modestly.[33] The effects of the other control variables remain similar to the earlier estimates. While these results are modest, they do suggest a link between price dispersion and internet usage. In particular, since internet fares are lower and since the fare differential shrinks with increased online purchases, price dispersion should fall. More generally, as the percentage of shoppers rises, dispersion may rise, fall, or stay the same. The issue of price dispersion is investigated more fully below.

To summarize, Table 4 provides three important results. First, controlling for ticket characteristics, capacity, route and carrier specific effects, online consumers on average pay roughly 13 percent less than the offline consumers. Second, equilibrium prices fall as the share of internet shoppers on a route rises. Finally, the results suggest that the potential savings of internet purchase in a lower internet usage market is marginally higher than the savings in higher internet usage markets.

---

[33] Though this effect may be surprising, it tends to provide a rationale for the decreased price dispersion. Increase in the share of internet usage implies more people buying on the internet such that a larger fraction faces similar prices such that the difference in the price paid on average will decrease. This is consistent with the theoretical prediction that as more informed consumers crowd a market, price dispersion will fall.

Effects of Internet on Price Dispersion

Stahl (1989) provides evidence that price dispersion is not monotonic with respect to the percentage of informed consumers or searchers. Instead, when there are no searchers there is a degenerate distribution of prices at the monopoly price. Dispersion then rises as some consumers becomes informed, eventually reaching a peak before descending back to zero at the competitive level when all consumers are informed. Brown and Goolsbee (2002) find evidence of such a non-monotonic relationship between online search and offline price dispersion for term life insurance. They attribute much of the non-linearity to the gradual maturation of internet markets.[34] They argue, based on Stahl(1989), that starting with a small enough share of internet users, giving a small group of customers access to more information will initially cause dispersion to rise but as more and more customers become informed, dispersion will eventually fall. Starting with a sufficiently high fraction of searchers (high internet share usage), Stahl's model would predict a monotonic negative relationship between internet purchases and price dispersion. We investigate similar issues but evaluate differences in price dispersion in the cross-section, exploiting differences in internet usage.

Following Brown and Goolsbee (2002) we adopt a two step approach to investigate the relationship between price dispersion and internet share. First, we estimate an hedonic pricing model, and second we examine the residuals of that model to assess variation in price dispersion. The first equation is:

[34] See Bailey (1998) and Brynjolfsson and Smith (1999) for implications of mature and immature internet markets.

$$\ln(\mathrm{Price}_{ijtk})=\beta_0 + \beta_1 (R_{ijtk}) + \beta_2 (\mathrm{Loadfactor}_{ijtk}) + \beta_3 (\mathrm{Route}_j) + \beta_4 C_{ijtk} + \beta_5 (\mathrm{DEP}_{ijtk}) + \beta_6 (\mathrm{RET}_{ijtk}) + \varepsilon_{ijtk} \qquad (2)$$

where inclusion of Route_j means we now include route fixed effects.[35]

From (2) we obtain the residuals and calculate different measures of dispersion (standard deviation, range and coefficient of variation) of the residuals by route-carrier for each departure date. These residuals parallel those of Brown and Goolsbee who measure residuals in life insurance after taking into account policy characteristics. These residuals represent unexplained price dispersion within a carrier-route after taking into account the observable characteristics of tickets. To assess the factors affecting price dispersion, we regress various measures of dispersion on the share of tickets purchased on the internet and on market structure variables (Table A2).

Figure 6 presents a kernel regression between the standard deviation of the residuals and the share of internet usage.

This regression indicates that initially there is a reasonably stable level of price dispersion, but then as online purchases rise, dispersion drops significantly. Hence, an increase in internet share appears to have a significant effect on dispersion.

Table 5 presents a regression analysis of the relationship between price dispersion, internet share, and market structure variables.

[35] Refer to Table A1 for the estimated results of this equation.

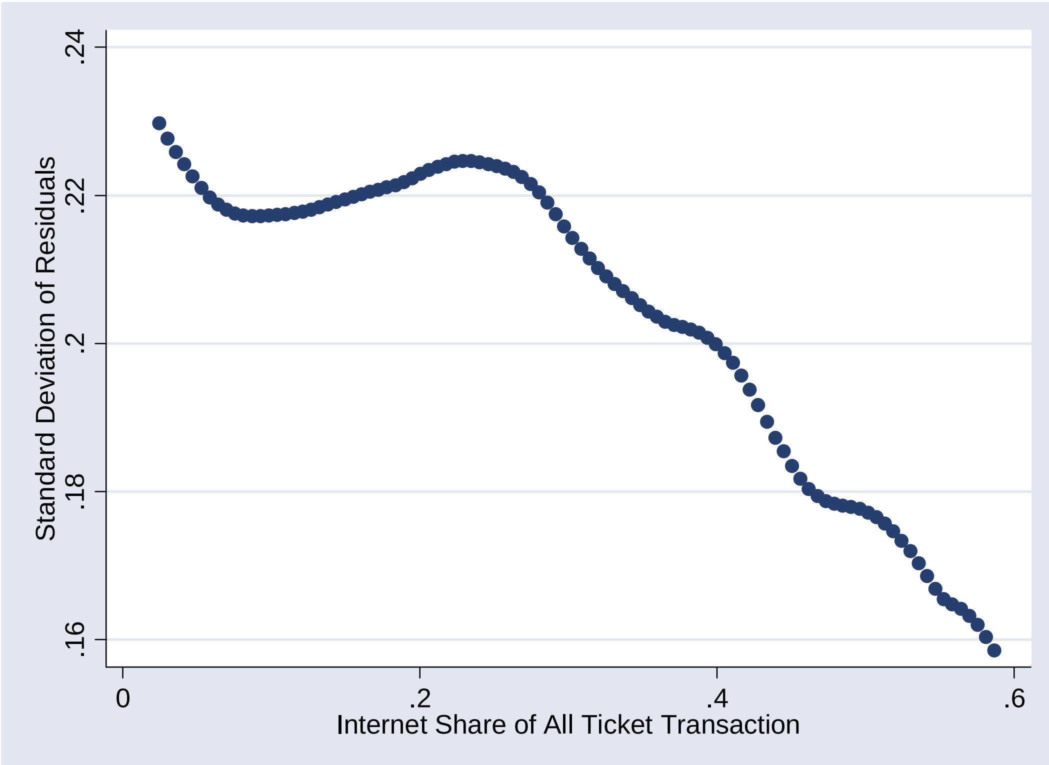


Figure 6
Price Dispersion and Share of Internet Purchases

Table 5
Regression of the Standard Deviation of Residuals on Market Structure Variables and Internet Usage

	(1) Standard deviation of residuals	(2) Standard deviation of residuals	(3) Standard deviation of residuals	(4) Standard deviation of residuals
Market share	0.0744	0.0633	0.0739	-0.0052
	20.13)**	(8.69)**	(19.97)**	(0.6)
HHI	-0.0946	-0.06	-0.0926	0.0133
	17.84)**	(5.98)**	(17.43)**	(1.24)
Internet Share	-0.0901	-0.0844	-0.0772	-0.0408
	(12.00)**	(11.10)**	(1.08)	(4.78)**
(Internet share)^2			0.2733	
			(0.98)	
(Internet share)^3			-0.6566	
			(2.01)*	
Carrier Fixed Effects (American Airlines omitted):				
Continental				-0.0042
				(1.24)
Delta				-0.0438
				(16.34)**
Northwest				-0.0157
				(4.63)**
United				0.0068
				(2.85)**
US Air				-0.0335
				(9.81)**
Frontier				-0.0428
				(10.60)**
Alaska				-0.0363
				(6.23)**
Hawaiian Airlines				-0.0716
				(3.34)**
America Mid West				-0.0272
				(5.48)**
American Trans Air				0.0009
				(0.07)
Midwest Express				-0.0662
				(4.22)**
Air Tran				-0.1835
				(21.67)**
Spirit				-0.0948
				(16.47)**
Sun County				-0.1083
				(12.62)**
Other Controls:				
Departure Date Fixed effects	No		No	Yes

Table 5, Continued

	(1) Standard deviation of residuals	(2) Standard deviation of residuals	(3) Standard deviation of residuals	(4) Standard deviation of residuals
Constant	0.267		0.259	0.218
	(88.61)**		(44.34)**	(24.18)**
Observations	20213		20213	20213
R-squared	0.030		0.030	0.090

Note: Absolute value of t statistics in parentheses * significant at 5%; ** significant at 1%; Source: Please refer to Table 1.

In column (1) of Table 5, we regress the standard deviation of the residuals on the share of internet usage, carrier market share, and the Herfindahl index of the route, without correcting for the endogeneity issues of the latter two variables.[36] In column (2) we run the same specification as in column (1) but we now correct for the potential endogeneity of market share and the Herfindahl index. The results show that price dispersion falls as the share of internet usage increases.

The relationship between price dispersion, market concentration and a carrier's share on a route parallels the results of Borenstein and Rose (1994). Using Gini coefficient as the measure of dispersion and without controlling for ticket characteristics, they find a 14 percent increase in dispersion from a one standard deviation (about 43 percent) decrease in the Herfindahl index. Their lack of control for ticket characteristics means that they measure dispersion with error, which could lead to poor estimation of

[36] We replicated these regressions with the range and coefficient of variation of the residuals as the dependent variables. The negative relationship between dispersion and share of internet usage was robust in all specifications irrespective of the choice of the dependent variable.

the effect of market concentration on the dispersion of prices. This possibility is significant because our results show the important role of ticket characteristics on prices.

We can use the results from Table 5 to investigate if the results of Borenstein and Rose (1994) vary when one controls for ticket characteristics. Using Table 5, a one standard deviation decrease in the market Herfindahl (about 53 percent) from the mean, leads to an increase in the standard deviation in prices of about 5 percent. These results fully control for observed ticket characteristics. Hence our results for the Herfindahl are consistent with the findings of Borenstein and Rose (1994) though the effects are diminished, owing either to our use of the standard deviation as a measure of dispersion or our control for observable ticket characteristics.[37]

Turning to market share, Table 5 shows that a one standard deviation of a carrier's market share (about 54 percent) from its mean increases the standard deviation in prices by about 4 percent. This effect is much stronger when compared to Borenstein and Rose (1994) who find a positive but insignificant effect of increased carrier's share on dispersion. Hence our results are broadly consistent with those of Borenstein and Rose.

The results suggest that as the share of internet usage rises, price dispersion falls. To check for non-linearity between dispersion and the share of internet usage we include the square and the cube of the internet usage share. Column (3) does not support a non-linear relationship consistent with the theoretical predictions of Stahl (1989) and the empirical results of Brown and Goolsbee (2002). In column (4) we include the carrier

[37] Note, the use of standard deviation of residuals to measure dispersion is to parallel the analysis of Brown and Goolsbee (2002).

specific effects and the carrier-departure day of week fixed effects respectively in addition to the variables used in column (2), correcting for the potential endogeneity in the market structure variables. This too suggests a negative relationship between dispersion and internet share usage. More importantly, our result suggesting an inverse relationship between price dispersion and share of internet usage is robust to these alternative specifications. Thus, we can conclude that rise in the share of internet purchases on a route decreases the overall dispersion in the fares.

Robustness Check

Table A3 of the appendix A replicates the regressions in Table 4 but uses a 5 percent match criterion. Using a 5 percent criterion increases the sample size by roughly 50 percent. The direct and indirect effects of the internet are stronger as compared to the main analysis. The results suggest that online prices on average are 15 percent less than offline prices while for every ten percent increase in the share of internet usage decreases the average prices by an additional 6 percent. The results are also consistent with our findings in the main analysis that online consumers in low internet usage markets save marginally more than in markets with a higher share of internet usage. These measures of the internet effects are very similar to those in the main analysis where the online prices are 13 percent lower than offline prices and a ten percent increase in internet share reduces the average prices by 5 percent.

The effects of the other variables are qualitatively similar, except that the coefficients for the maximum stay requirement switches signs and are significant. Also, the OLS estimates of the Herfindahl index switch signs to become positive and significant while the instrument variable regression estimates continue to be significantly negative and comparable to those in the main analysis. The rest of the results are qualitatively similar to those in the main analysis.

Implications of Consumer Surplus

In this section we provide some estimates of the overall magnitude of the savings the internet generates for customers. The analysis in Table 4 holds ticket characteristics constant. The internet, however, also helps customers identify the restrictions they are willing to accept by providing better information about the price/restriction tradeoff. Hence one can exclude the ticket characteristic regressors to identify the overall savings experienced by customers who buy on the internet. This approach likely overstates internet savings because it ignores potential selection bias in that internet customers may have found cheaper fares offline if the internet were unavailable. Still, such an approach does measure the overall savings, providing an upper bound on the gains consumers obtain from the internet. The results also show the dramatic impact of controlling for ticket characteristics when investigating price savings.

Column (1) of Table 6 presents a simple regression of the ticket prices and the online dummy variable along with return and departure day of the week, carrier and route fixed effects. We run a similar specification without the departure and return day of the week in column (2) of Table 6. Both specifications suggest that online consumers on average save about 44 percent on ticket prices as compared to offline consumers.[38] This result indicates that an online consumer looking for a low price ticket and willing to accept restrictions saves very substantial sums compared to customers buying through offline outlets.

Table 6 also provides additional analysis to investigate whether our matching protocol has influenced estimated savings on the internet. Columns 3 and 4 of Table 6 present regressions estimating internet savings using both matched and unmatched observations so that these results can be compared to the results of Columns 1 and 2, which include only matched observations. Inspection of the results shows that when one includes the unmatched observations, estimated internet savings rise from about forty-four percent to about fifty-one percent. Hence the results suggest that omission of the unmatched observations may lead to a modest underestimation of the price savings from the internet.

[38] Note that this result is driven in part by the adverse selection problem where customers who want lower fares are willing to accept various restrictions to get them and are more prone to shop on the internet. We ran a simple logit model with online dummy variable as our dependent variable and the ticket characteristics as the explanatory variables. We did find strong evidence that ticket characteristics associated with lower prices, namely Saturday stay-over, travel restriction, travel during non-peak times, and advance purchase requirements are more likely to be bought online than offline. However, tickets associated with minimum and maximum stay restrictions were found to be less likely to be bought on the internet.

Table 6
Benefit from Purchasing on the Internet

	(1) log(Fare)	(2) log(Fare)	(3) log(Fare)	(4) log(Fare)
	Columns 1 & 2 Include Only Matched Observations		Columns 3 & 4 Include Matched & Unmatched Observations	
Online	-0.414	-0.442	-0.5101	-0.5501
	(176.86)**	(183.56)**	(422.79)**	(457.58)**
Departure Day of Week (Sunday omitted):				
Monday	0.001		-0.0061	
	(0.250)		(3.93)**	
Tuesday	-0.015		-0.0309	
	(5.21)**		(19.52)**	
Wednesday	-0.035		-0.0529	
	(12.06)**		(32.38)**	
Thursday	-0.080		-0.0943	
	(26.74)**		(56.70)**	
Friday	-0.147		-0.1648	
	(47.22)**		(95.82)**	
Saturday	-0.116		-0.2033	
	(31.12)**		(100.73)**	
Return Day of the Week (Sunday omitted):				
Monday	-0.392		-0.3517	
	(131.94)**		(226.79)**	
Tuesday	-0.356		-0.3092	
	(130.92)**		(208.61)**	
Wednesday	-0.362		-0.3116	
	(143.75)**		(223.23)**	
Thursday	-0.348		-0.2726	
	(145.60)**		(205.81)**	
Friday	-0.356		-0.2667	
	(155.85)**		(208.58)**	
Saturday	-0.411		-0.3501	
	(113.81)**		(182.90)**	
Carrier fixed effects	Yes	Yes	Yes	Yes
Route fixed effects	Yes	Yes	Yes	Yes
Constant	5.965	5.725	5.833	5.576
	(629.57)**	(593.45)**	(1055.28)**	(1006.94)**
Observations	523618	523618	1850991	1850991
R-squared	0.370	0.300	0.380	0.340

Note: Absolute value of t statistics in parentheses; * significant at 5%; ** significant at 1%;
Source: Please refer to Table 1.

Conclusions

This paper has provided an in-depth empirical analysis of the effects of the internet on the price of airline tickets. The research makes use of the most complete data set ever used to analyze airline pricing. This novel data set includes data on individual ticket transactions including ticket characteristics, carrier, flight level load factor at purchase, measures of peak departure and return times, the date of issue, other hedonic factors affecting prices, and an indicator to denote if the ticket was purchased online. This paper is one of only a few studies of internet pricing using data for contemporaneous online and offline transactions in comparable geographic markets.

While controlling for numerous observed ticket characteristics, carrier and route effects the results show that online prices are about 13 percent less than the offline prices. The analysis also shows that a ten percent increase in the share of tickets sold online decreases average prices by 5 percent, with more of its effect coming in the form of lower offline prices. The paper also finds evidence that increased online shares decrease price dispersion.

The paper has also used these new data to present a more complete analysis regarding the impact of market structure on the level and dispersion of airline prices. The results largely confirm the finding of Borenstein and Rose. Even when one controls for ticket characteristics and peaking variables, the impact of market share and market concentration on the level and dispersion of ticket prices is qualitatively similar to that found by Borenstein and Rose.

CHAPTER IV

HETEROGENEITY IN CONSUMERS AND CHANNELS OF SALE: IMPLICATIONS FOR PRICE AND PRICE DISPERSION

Introduction

The effects of the internet on the levels and dispersion in prices as analyzed in Chapter III rests on the implicit assumes that all internet consumers are searchers. This assumption is at par with the assumptions made in the existing literature comparing market efficiencies between internet and non-internet markets. This assumption of all internet consumers are searchers fails to accommodate the role of internet as a convenient medium of transaction – less transaction time and ease of shopping. These consumers using internet as a convenient transaction medium are not traditional 'shoppers', as the present literature does.

This chapter presents an alternate model of consumer search which accommodates both searchers and non-searchers in each of the electronic and traditional markets. Using this consumer heterogeneity and randomized pricing strategy by firms, we derive implications of increased internet usage on the average prices paid by both the searchers and non-searchers in the overall market for a homogenous good. This study also investigates the relationship between price dispersion and internet usage given the new consumer dichotomy in the overall market where the market overall is a linear combination of the electronic and traditional market.

There exists varied rationale for price dispersion to be an equilibrium phenomenon. Theoretical models have tried to rationalize the phenomenon of price dispersion associated with a homogenous good by invoking ex-ante consumer heterogeneity with respect to search costs (Shilony, 1977); information availability (Stiglitz and Salop, 1977); the propensity to search (Wilde and Schwartz, 1979) or ex-ante producer heterogeneity arising from differences in production costs (Reinganum, 1979); and marketing strategies (Varian, 1980).

Burdett and Judd (1983) argue that the critical aspect of a model that leads to equilibrium price dispersion is the existence of ex-post heterogeneity in consumer information that can arise with or without ex-ante heterogeneities among the consumers and/or the producers.

Equilibrium price dispersion can probably be most intuitively rationalized by the presence of a positive marginal cost of obtaining each price quote. This serves as a plausible rationale of price dispersion in both conventional (Pratt et al (1979), Carlson and Pescatrice (1980), Sorensen (2000)) as well as for some virtual markets (Smith, Bailey and Brynjolfsson (1999), Bakos (2001)). The marginal cost of search in conventional markets involves visiting an additional store, while in the online market it involves visiting an additional website if no price comparison site exists, or visiting the website of the actual vendor if the price comparison site only involves the listing of the prices or `deals' of the different vendors on their webpage (example Kelkoo.com (Baye, Morgan and Scholten (2004)). The present literature in electronic commerce, however, assumes a zero marginal cost of obtaining an additional quote for all internet consumers.

This paper specifically addresses this shortcoming in the literature and proposes an alternate model of search in the electronic markets.

Stahl (1989) rationalizes equilibrium price dispersion using a sequential search model and randomized pricing strategy by firms. Stahl (1989) assumes a market with two groups of consumers – a fraction of shoppers who face a zero search cost and another fraction of non-shoppers who face a positive cost of obtaining additional price quote beyond the first quote. The author argues that under search a framework, price dispersion is an equilibrium outcome. Stahl (1989) posits a non-monotonic relationship between the fraction of searchers and dispersion with the dispersion initially increasing and then monotonically decreasing as the share of informed consumers rises.

Price dispersion is not only a widespread phenomenon in the physical (offline) market but in the internet (online or electronic) market as well. The drivers of price dispersion, however, may differ in the two markets. Brynjolfsson and Smith (1999) argue that much of the dispersion in the internet market is attributed to retailer heterogeneity with respect to branding, awareness and trust. The discussion with respect to these drivers of dispersion, unfortunately, is beyond the scope of this research.

The existing literature in electronic markets implicitly assumes all consumers in the internet market to be shoppers, searching for the lowest price. The internet, however serves a dual role. Firstly, it provides the consumers with a platform to search exhaustively and at a lower cost. Secondly, it serves as a channel for convenient mode of transaction for the more convenience shoppers. Most of the models fail to accommodate the convenience role of the internet. Existing research in electronic markets implicitly assume the internet consumers to be intrinsically shoppers, those looking for the lowest price. This assumption incorrectly categorizes those customers as shoppers who use the internet only as a convenient medium of transaction. Convenience of shopping on the internet mainly constitute of less transaction time, less effort and hassle free shopping experience. Additionally, the literature also assumes consumers in the traditional markets to be lesser price sensitive and categorize them to be non-searchers. More generally, there exist a group of consumers who searches extensively (at least, all stores in a region) to find the lowest price available or the 'best deal', making them searchers as opposed to non-searchers in the more traditional markets.

This paper builds on these motivations to include both searchers and non-searchers in each of the online and offline market, which otherwise has been assumed away in the existing theoretical literature[39] and measure its implications on the average level of prices and dispersion.

I build on the basic framework of Stahl (1989) to include the group of more and less informed searchers (consumers) in each market and derive its implications for the average price and price dispersion in the overall market, where the overall market is the linear combination of the offline and online markets.

---

[39] For example, consider the market for airline tickets. Airline tickets can be bought either by calling on or visiting the different travel agents in a certain neighborhood (offline) or through the different travel websites (online). Among the consumers who buy their tickets from the travel agents, there exist a fraction of consumers who would call each and every travel agent in the nearby geographical area to get the best possible 'deal'. This group can be referred to as the high intensive searchers in the offline market. On the other hand, the business travelers would take no such incentive to find the lowest price and would buy from any travel agent, in most cases a pre-assigned agent. Similarly, in the online market, there will exist a fraction of online consumers who would just visit one single website (maybe, due to ease of navigating the website, prior experience on this website, "branding effect") and purchase their ticket without searching for the lowest price that may be available in the other travel websites, thus constituting our group of low intensive searchers in the online market.

For simplicity, consider a market for a homogenous good which consists of two identical but independent sub-markets -- a conventional market (offline) and an internet market (online). The firms in each of these sub-markets compete within their own periphery but not with those outside their competitive realm. In other words, we assume a state of the world where the offline (physical) sellers compete with the other offline sellers in the market, while the online sellers compete only among themselves. Thus, we assume that there is no competition between the online and offline sellers in the market.[40]

We show that the lower search costs in the online market lowers the expected price to be paid by all groups of consumers in the overall market, but only at the consequence of more dispersed prices, than otherwise. This paper, thus contributes to the existing theoretical literature by providing a theoretical framework that includes both searchers and non-searchers in each individual market and derives implications of improves information on average level of prices and dispersion.

Theoretical Framework

Assumptions

We assume for simplicity, that a fraction (λ) of the consumers of the homogenous good is online consumers while the remaining ($1-\lambda$) fraction of the consumers is

[40] For example, in the market for airline tickets, the airlines sell tickets either through the physical travel agents or through the web-sites. The physical travel agents will compete for the consumers in the same geographical area but is highly unlikely to compete with any of the travel websites. In this sense, the offline and online travel agents compete in their own market periphery but not with those outside their realm.

restricted to the offline market, with the population mass being normalized to one. Each consumer has a reservation price[41] and will buy the good if and only if the price offered is no more than her reservation price.

In each of these sub-markets (online and offline) we further assume that the consumers are either `high intensity searchers' or `low intensity searchers'. I define the `high intensity searchers' as those who have a very low (for simplicity I assume it to be zero) marginal cost of visiting an additional store or website, beyond their first price quote. Hence by our definition, the high intensity searchers are expected to search each and every store (website) and choose the lowest price, provided the lowest price is no more than their reservation price. On the other hand, the `low intensity searchers' are faced with a high marginal cost of search beyond their first price quote that prevents them from searching extensively.[42] It may also be the case that the buying decision of low intensity searchers is not solely motivated by price but also by other attributes of the good.

Let μ_i represent the fraction of the high intensity searchers in the online market while $(1- \mu_i)$ be the fraction of low intensity searchers in the online market. Similarly, μ_o and $(1- \mu_o)$ represent the fraction of high and low intensity searchers in the offline market respectively. Finally, we assume that there is no cross search between the offline and online markets. In our model, we do not allow for the possibility that people will

[41] This can also be interpreted as his maximum willingness to pay for the good.
[42] One can think of it as these consumers being subject to a very high opportunity cost of time

first search in the offline market and then in the internet market or vice-versa. This assumption is essential in keeping the model simple and tractable.

Optimal Search Rule

Following Stahl (1989), we assume that each consumer knows the distribution of prices but gets to know the price offered at a particular store only by visiting it. We adopt three simple rules of search (only for non-shoppers). Firstly, a consumer will search the next store if the minimum price observed till now is strictly greater than his reservation price (R). Secondly, a consumer will stop search and buy from the store if the offered price is less than his reservation price (R). And finally, we assume that a consumer is indifferent between searching and not searching an additional store (website) if the offered price is exactly equal to his reservation price, R. In a symmetric Nash equilibrium (NE), a consumer samples from the same price distribution in each market, S(P) each time he visits a new store (web-site).

Given the shopping rule as in Stahl (1989), the indifference case needs some discussion. Though it does not matter for the positive search cost consumers ($c>0$), the case for the zero search cost ($c=0$) consumers does matter. If the high intensity searchers with zero search cost stop at the first store (web-site) with price less than or equal to reservation price ($P \leq R$)then the monopoly price is an unique NE since all consumers visit the store only once which is not reasonable given the heterogeneity in search cost. Thus we assume that the shoppers do not stop before they sample at least two stores and/or prices. However, it can be proved that with probability one the zero search cost shoppers will shop all the n stores and buy from the store that offers them the lowest price.[43]

Store Price Setting Behavior

The process of sellers setting a price is a two stage process. In the first stage we assume an exogenous consumer reservation price r, and we determine a Nash equilibrium (NE) in price-setting conditional on r.

Let F(p;r)denote the cumulative distribution function of prices adopted by each of the n stores in such a conditional NE. This distribution will generate a reservation price R. In the second stage we want a consistent reservation price r* that will produce a NE distribution F(p;r*) which in turn will produce R=r* The following provides us with:

Lemma 1: Given μ€(0,1), if F(p,r) is a NE distribution conditional on reservation price r, then it is atomless.

The intuitive argument is that the profits could be increased by undercutting, so price distributions with atoms are not optimal. Let p^h be the maximal element in the support of F(p,r) and let p_l be the minimal element.

Lemma 2: If F(p;r) is an NE distribution conditional on reservation price r, then $p^h = \min\{r, p^m\}$.

To state in simple terms the upper bound of the support of F(p;r) is the minimum of the reservation and the monopoly price, p^m. Say if store j's price $p_j>p_l$ then none of the shoppers (c=0) will buy and only the c>0 consumers will but at $p_j=p^h<r$. Also pricing above p^m is not optimal as p^m is the monopoly profit maximizing price. But

[43] Please refer to Stahl (1989) for a complete proof.

again $p_j = p^h < r$ yields a lesser revenue than $p_j = r$. Hence, given a conditional NE distribution F(p;r)consumers with positive search costs (c>0) will always observe $p_j \leq r$ and hence will stop with probability 1 at the first stop they visit. Only the true searchers (c=0) will indulge in real search activity.

Using Lemma 1 and 2 we can say that F(p;r) is a conditional NE if and only if the expected profits is equal for all p in the support of F(p;r). Also F(p;r) is a consistent NE if F is a conditional NE and r is a consistent reservation price.

Then

$$\pi_i(\mathrm{P}) = \pi_i(\mathrm{r})$$

or, $P(\mu + \frac{(1-\mu)}{n})(1-F(P))^{n-1} + P(\frac{(1-\mu)}{n})(1-(1-F(P))^{n-1} = r(\frac{(1-\mu)}{n})$, which yields

$$(1)\ F(P) = 1 - \left(\frac{(1-\mu)(r-P)}{nP\mu}\right)^{\frac{1}{n-1}}.$$

One can calculate the lower bound p_l by setting F(P)=0 as shown in the next section.

Finally, we establish that the consumer reservation price r is consistent with F(P;r) or there exists at least one such consistent r. Now R was the unique solution to EB(z)=c, if one exists else R=∞. Thus a consistent reservation price $r^* \leq p^m$ must satisfy

$$g(r) = \int_{p_l}^{r^*} D(p)F(p, r^*, \mu, n)dP - c = 0$$

$$\Rightarrow \int_{p_l}^{r^*} D(p)F(p, r^*, \mu, n)dP = c$$

$$\Rightarrow \int_{p_l}^{r^*} F(p, r^*, \mu, n)dP = c$$

$$\Rightarrow \int_{p_l}^{r^*} \{1 - \left(\frac{(1-\mu)(r-P)}{nP\mu}\right)^{\frac{1}{n-1}}\}dP = c$$

$$\Rightarrow g(r) = (r^* - P_l) - [\frac{(1-\mu)}{n.\mu}]^{\frac{1}{n-1}} \int_{p_l}^{r^*} [\frac{(r-P)}{P}]^{\frac{1}{n-1}} dP = c$$

Deriving the Distribution of Prices

Given the fraction of high and low intensive searchers, reservation price and search cost in each market, we now try to derive the price distribution functions in each of these markets individually.

Since, we have assumed at the offset that the two markets are identical, hence, the deriving the price distribution of one market will pave the way for the other.

We have assumed, that μ_o represent the high intensity searchers in the offline market. Since, these consumers would search all the stores and buy from the lowest priced seller, the remaining (1- μ_o) low intensive searchers will now be distributed evenly among the n-sellers, where each seller has (1- μ_o)/n share of captive consumers.

Let, $P_{-i}^o = \{P_1^o, P_2^o, P_3^o, ..., P_{i-1}^o, P_{i+1}^o, ... P_n^o\}$, where, P^o denotes an offline price. The expected profit of firm i when it chooses price P is given by:

$$= P^o(\mu_o + \frac{(1-\mu_o)}{n})$$

$$\pi_i = P^o(\frac{\mu_o}{m} + \frac{(1-\mu_o)}{n}) \text{ if } P^o < \min(P^o_{-i})$$

$$= P^o \frac{(1-\mu_o)}{n}$$

There exists a symmetric equilibrium that involves mixed strategies where each firm prices according to a continuous cumulative density function $F(\mathrm{P^o})$ on $[\mathrm{P_o},\ \mathrm{r^o}]$ where $\mathrm{P_o}$ refers to the lower bound of the distribution of the offline (physical) market and $\mathrm{r^o}$ refers to the reservation price of the consumers in this market.

Now, the expected profits for all prices in the support of $F(\mathrm{P^o})$ should be equal, that is

$$\pi_i(\mathrm{P^o}) = \pi_i(\mathrm{r^o})$$

or, $\mathrm{P^o}(\mu_o + \frac{(1-\mu_o)}{n})(1-F(\mathrm{P^o}))^{n-1} + \mathrm{P^o}(\frac{(1-\mu_o)}{n})(1-(1-F(\mathrm{P^o}))^{n-1} = r(\frac{(1-\mu_o)}{n})$

$$(2)\ \ F(\mathrm{P^o}) = 1 - \left(\frac{(1-\mu_o)(r^o - \mathrm{P^o})}{n\mathrm{P^o}\mu_o}\right)^{\frac{1}{n-1}}$$

To derive the lower bound of $F(\mathrm{P^o})$ we set it equal to 0 to get:

$$(3)\ \ P_o = \frac{(1-\mu_o)r^o}{1+n\mu_o-\mu_o}$$

Similarly for the internet market we have:

$$(4)\ \ G(\mathrm{P^i}) = 1 - \left(\frac{(1-\mu_i)(r^i - \mathrm{P^i})}{n\mathrm{P^i}\mu_i}\right)^{\frac{1}{n-1}}$$

Also, $P_i = \frac{(1-\mu_i)r^i}{1+n\mu_i-\mu_i}$ where P_i represents the lower bound of the distribution of prices in the online market.

Deriving the Reservation Price

From the stopping rule, as discussed above, it is quite evident that a consumer will search an additional store as long as the marginal benefit (gain) from searching is at least greater or equal to the cost of searching the additional store. The expected gain from search in the offline market $g(r^o)$ is given by:

$$g(r^o) = \int_{P_o}^{r^o} (r^o - \mathrm{P^o})dF(\mathrm{P^o})$$

or, $g(r^o) = (P_o - r^o)F(P_o) + \int_{P_o}^{r^o} F(\mathrm{P^o})d\mathrm{P^o}$

Now F(P_o)=0

Hence, $g(r^o) = \int_{P_o}^{r^o} F(\mathrm{P^o})d\mathrm{P^o}$

Assume that, for every additional store that a consumer visits after the first visit, incurs a cost of c_o.Also, the expected gain is equal to the marginal search cost and can expressed as:

$$g(r^o) = c_o$$

Simplifying the expression for g(r^o) we get:

(5) $g(r^o) = (r^o - P_o) - \left(\frac{1-\mu_o}{n\mu_o}\right)^{\frac{1}{n-1}} \int_{P_o}^{r^o} [\frac{(r^o - \mathrm{P^o})}{\mathrm{P^o}}]^{\frac{1}{n-1}} d\mathrm{P^o}$

The above equation does not have a closed form solution and differs with the value of n. In the simplest duopoly case, with n=2 we get:

$$(6)\ r^o = \frac{c_o}{1-(\frac{1-\mu_o}{2.\mu_o})\ln\left(\frac{1+\mu_o}{1-\mu_o}\right)}$$

The reservation price in the internet market can be derived in a similar manner as the offline market. For the internet market, we assume that the consumers incur a search cost for every additional website they search beyond their first visit. However, we assume that the search cost for the internet consumers is a function of the offline search cost given by $c_i=\delta c_o$, where $0<\delta<1$. Following the above approach, we yield similar conditions in the internet market given by:

$$(7)\ g(\mathrm{r^i}) = (\mathrm{r^i} - \mathrm{P^i}) - (\frac{1-\mu_i}{n.\mu_i})\int_{P_i}^{\mathrm{r^i}}\left[\frac{\mathrm{r^i}-\mathrm{P^i}}{\mathrm{P^i}}\right]^{\frac{1}{n-1}}.d\mathrm{P^i}\text{, and}$$

$$(8)\ \mathrm{r^i} = \frac{\delta c_o}{1-(\frac{1-\mu_i}{2.\mu_i})\ln(\frac{1+\mu_i}{1-\mu_i})}$$, where n=2 and $\mathrm{r^i}$ represents the reservation price in the internet market.

Distribution of Prices in Individual Markets

In this section we derive the expected prices in each of the individual markets. The consumers know the distribution of the prices faced by them in the internet or the offline market as the case maybe, and given the known distribution of prices, the consumers can calculate the expected price they are likely to face.

We begin with the offline market. The distribution of prices in the offline market, as derived earlier, is given by

$$F(\mathrm{P^o}) = 1 - \left(\frac{(1-\mu_o)(r^o - \mathrm{P^o})}{n\mathrm{P^o}\mu_o} \right)^{\frac{1}{n-1}}$$

Therefore, $f(\mathrm{P^o}) = \dfrac{\partial F(\mathrm{P^o})}{\partial(\mathrm{P^o})} = \dfrac{r^o}{\mathrm{P^o}(r^o - \mathrm{P^o})} \dfrac{1}{(n-1)}[1 - F(\mathrm{P^o})]$

Now, $E(\mathrm{P^o}) = \int_0^\infty [1 - F(\mathrm{P^o})] d\mathrm{P^o}$, which can be alternatively written as

$$(9)\;\; E(\mathrm{P^o}) = \mathrm{P_o} + \int_{\mathrm{P_o}}^{\mathrm{r^o}} \left[\frac{(1-\mu_o)(\mathrm{r^o} - \mathrm{P^o})}{n.\mathrm{P^o}.\mu_o} \right] d\mathrm{P^o}$$

Similarly, we can show that the expected price in the internet market will be given by

$$(10)\;\; E(\mathrm{P^i}) = \mathrm{P_i} + \int_{\mathrm{P_i}}^{\mathrm{r^i}} \left[\frac{(1-\mu_i)(\mathrm{r^i} - \mathrm{P^i})}{n.\mathrm{P^i}.\mu_i} \right] d\mathrm{P^i}$$

Proposition 1: The expected price decreases as the number of high intensity searchers in each of the individual market rises.

Proof: See Appendix B.

Price Distribution in the Overall Market

By assumption, the entire consumer population is divided into two groups depending on their accessibility to the internet for buying and searching behavior for airline tickets. We assume that λ fraction of the total consumer population has access to internet while the remaining (1-λ) fraction do not. The price distribution in the offline and internet market are represented by $F(P^o)$ and $G(P^i)$ respectively. Given this, the overall market distribution for the airline ticket prices will be a weighted sum of the distributions in the individual markets where the weights are the fraction of the population having access to the internet and the fraction not having the same, which is given by the equation

(11) $M(P) = \lambda.G(\mathrm{P}^{\mathrm{i}}) + (1-\lambda).F(\mathrm{P}^{\mathrm{o}})$

As λ→1, the overall market distribution of the prices M(P) will collapse to the internet distribution of the prices G(P^{i}). Similarly as λ→0, the overall market distribution of the prices M(P) will collapse to the internet distribution of the prices F(P^{o}). This property helps us to justify our above expression of the overall market distribution of prices as an weighted sum of the distributions in each of the individual markets.

The expected price in the overall market, consisting of both the online and offline markets can be written as

(12) $E[M(P)] = \lambda.E[G(\mathrm{P}^{\mathrm{i}})] + (1-\lambda).E[F(\mathrm{P}^{\mathrm{o}})]$

The expected price in each of the individual market however will depend on the shares of searchers and the non-searchers in each market. Hence, the above equation for the expected price in overall market can be modified as

(13) $E[M(P)] = \lambda\{\mu_i.E(\mathrm{P}^{\mathrm{i}}{}_{\min}) + (1-\mu_i)E(\mathrm{P}^{\mathrm{i}})\} + (1-\lambda)\{\mu_o.E(\mathrm{P}^{\mathrm{o}}{}_{\min}) + (1-\mu_o)E(\mathrm{P}^{\mathrm{o}})\}$

or, $E[M(P)] = \lambda.B(\mu_i) + (1-\lambda)B(\mu_o)$

where $B(\mu_i) = \mu_i.E(\mathrm{P}^{\mathrm{i}}{}_{\min}) + (1-\mu_i)E(\mathrm{P}^{\mathrm{i}})$

and $B(\mu_o) = \mu_o.E(\mathrm{P}^{\mathrm{o}}{}_{\min}) + (1-\mu_o)E(\mathrm{P}^{\mathrm{o}})$

Thus, $\dfrac{\partial E[M(P)]}{\partial \lambda} = B(\mu_i) - B(\mu_o)$

$$= \frac{\partial E[M(P)]}{\partial \lambda} = [\mu_i . E(\mathrm{P}^{\mathrm{i}}{}_{\min}) - \mu_o . E(\mathrm{P}^{\mathrm{o}}{}_{\min})] + [(1-\mu_i)E(\mathrm{P}^{\mathrm{i}}) - (1-\mu_o)E(\mathrm{P}^{\mathrm{o}})]$$

or, $\frac{\partial E[M(P)]}{\partial \lambda} < 0$ if $\mu_o = \mu_i$

Now, $\mathrm{var}[M(P)] = E[M(\mathrm{P}^2)] - E[M(P)]^2$

It can be shown that, $\frac{\partial \mathrm{var}[M(P)]}{\partial \lambda} \lesseqgtr 0$ followed by $\frac{\partial^2 \mathrm{var}[M(P)]}{\partial \lambda^2} < 0$ [44]

Proposition2: The expected price in the overall market falls as share of internet access rises. In other words, the expected price to be paid by both the high and low intensive searchers decreases as the share of internet access increases.

Proof: See Appendix B.

Corollary 2.1: For a duopoly, the expected price to be paid by the high intensive searchers in the internet market will be lower than their offline counterpart.

Proof: See Appendix B.

Corollary 2.2: For a duopoly, the expected price to be paid by the low intensive searchers in the internet market will be lower than their offline counterpart.

Proof: See Appendix B.

Proposition 3: The price dispersion in the overall market initially increases as the share of internet access rises, but decreases eventually with further increase in the share of internet access.

Proof: See Appendix B.

[44] For complete proof, please refer to Appendix B.

Conclusion

The existing search cost models and their applications to internet markets implicitly assume the consumers in the online markets to be essentially the searchers while those in the traditional markets typically are assumed to constitute the non-searcher population. Internet typically serves a dual role. Firstly, it provides the consumers with a platform to search exhaustively and at a lower cost. Secondly, it serves as a channel for convenient mode of transaction for the more convenience shoppers. The existing literature has primarily focused on the low search cost on the internet while undermining the convenience property of the internet. This paper provides an alternate model of search that includes a group of both searchers and non-searchers in each of the traditional and internet markets respectively for a homogenous product. The study then exploits the search cost based consumer heterogeneity with randomized pricing strategies by firms to evaluate the implication of increased internet usage on average price and dispersion in the overall market.

The results suggest that as the share of people with internet access rises, it creates a positive externality by reducing the expected price to be paid by both the low and high intensive searchers in the overall market. The study also shows that the high intensive searchers in the internet market pay a lower price than the high intensive searchers in the offline market. The results are similar for the low intensive searchers in the two markets. Finally, the results also suggest a non-monotonic relationship between increased internet usage and dispersion, with the dispersion initially increasing followed by a monotonic decline as the fraction of people with access to internet rises.

CHAPTER V

CONCLUSION AND SCOPE OF FUTURE RESEARCH

This dissertation analyzes the economic implications of increased internet usage on average prices and price dispersion, both theoretically and empirically. There exists a broad array of literature analyzing the effects of consumer search behavior and its effect on average prices and price dispersion. The existing theoretical literature offers a reasonable analysis of the economic implications of consumer search on prices and price dispersion. Recent literature in electronic commerce has adhered to the traditional search models in addressing research pertaining to electronic markets: how average prices are affected as more people search on the internet, how increased internet usage affects price dispersion in different markets, to mention a few. Most of these models, however, fall short of characterizing the electronic markets completely.

Internet serves a dual role. Firstly, it provides the consumers with a platform to search exhaustively and at a lower cost. Secondly, it serves as a channel for convenient mode of transaction for the more convenience shoppers. Most of the models fail to accommodate the convenience role of the internet.

Existing research in electronic markets implicitly assume the internet consumers to be intrinsically shoppers, those looking for the lowest price. This assumption incorrectly categorizes those customers as shoppers who use the internet only as a convenient medium of transaction. Convenience of shopping on the internet mainly constitute of less transaction time, lesser effort and stress free shopping experience. Additionally, the literature also assumes consumers in the traditional markets to be lesser price sensitive and categorize them to be non-searchers. More generally, there exist a group of consumers who searches extensively (at least, all stores in a region) to find the lowest price available or the ‘best deal’, making them searchers as opposed to non-searchers in the more traditional markets.

Chapter IV of this dissertation presents a model that accommodates both groups of searchers and non-searchers in the electronic and traditional marketplace. I then use this modified theoretical framework to derive implications of the increased access to internet on average prices and price dispersion in the overall market for a homogenous good, where the overall market is a linear combination of the online and traditional channels of sale.

The model predicts that as share of internet usage increases the expected price in the overall market falls. The results suggests that given the same share of searchers in the two markets but a lower search cost in the internet market, the expected price to be paid by both searchers and non-searchers in the internet market will be lower than the expected price to be paid the consumers in the traditional markets. More generally, the model implies that the overall effect of the increased internet usage on the average price to be paid by the searchers and non-searchers in the two markets is dependent on the search cost. The model also provides evidence of a non-monotonic relationship between the proportion of internet usage and price dispersion, with the price dispersion initially increasing as share of internet usage but decreases as the proportion rises, consistent with Stahl (1989).

One possible shortcoming of this model is the empirical tractability of the underlying theoretical implications. One would need sufficiently rich data on consumer search behavior in both internet and physical markets that would allow researchers to classify consumers as searchers or non-searchers, with much precision. And clearly, this dissertation falls short, to put the theoretical hypotheses of this model to rigorous empirical testing.

Empirical testing of consumer search models demands some strict assumptions with respect to the definitions of searchers and non-searchers. Due to the ease of empirical tractability, it has become a commonplace assumption in internet literature to define the consumers in the electronic markets as 'searchers' and non-internet consumers to be 'non-searchers'. This dissertation adopts this standard assumption to study the empirical effects of the internet usage and their effects on the prices for airline seats in Chapter III.

This study provides the most comprehensive analysis of the impact of internet purchases on pricing for a single industry, using a unique data set from the airline industry. The overall goal of the paper is to investigate the effects of internet sales on prices paid for airline tickets. The data set includes variables reflecting numerous observable ticket characteristics that together determine the large majority of the variation in ticket prices. The analysis also controls for load factors at purchase, network peak times, market structure, and route characteristics. The analysis measures the direct impact of internet purchase for those customers who buy on the internet, controlling for ticket and market characteristics. In addition, the analysis investigates the effects of increased internet purchases on average prices for both those who buy online and those who buy offline. Finally the analysis examines the impact of internet purchases on the dispersion of prices for online fares, offline fares, and fares overall. The analysis is important in part because it permits an integrated analysis of these various factors in the context of a single industry using actual transactions data.

This study undertakes an in-depth empirical analysis of the effects of the internet on the price of airline tickets. The research makes use of the most complete data set ever used to analyze airline pricing. The study makes use of a novel data set that includes the data on individual ticket transactions including the ticket characteristics associated with each transaction, carrier, flight level load factor, the date of issue, departure and return date, indicator to denote if the ticket was purchased online or offline and other hedonic factors affecting the pricing of airline seats. This paper is only one of a few studies of internet pricing using data for contemporaneous online and offline transactions and where the ultimate seller is the same. While controlling for numerous observed ticket characteristics, carrier and route effects the results show that online prices are about 13 percent less than the offline prices. The analysis also shows that a ten percent increase in the share of tickets sold online decreases average prices by 5 percent, with more of its effect coming in the forms of lower offline prices. The study also finds evidence that increased online shares decrease price dispersion.

In sum, this dissertation provides an in-depth analysis of the economics of electronic markets. It provides one of the most comprehensive empirical analysis of the effects of internet on average prices and price dispersion in a major industry (airlines) using a unique contemporaneous online and offline transaction data, which the contemporary empirical literature is significantly void of. The dissertation also provides a theoretical framework that accommodates both searchers and non-searchers in internet and non-internet markets and analyzes their implications on price levels and dispersion. This dissertation falls short of empirical testing the predictions of the theoretical model attributed to the lack of adequate data necessary and leaves this task for future research.

REFERENCES

Bailey, J., 1998, ‘Intermediation and Electronic Markets, Aggregation and Pricing in Internet Commerce,’ Ph.D. Dissertation*, Technology, Management and Policy, Massachusetts Institute of Technology, Cambridge, MA.*

Bakos, Y., 1997, ‘Reducing Buyer Search Costs, Implications for Electronic Marketplaces,’ *Management Science*, 43 (12), pp.1676-1692.

Bakos, Y., 2001, ‘The Emerging Landscape for Retail E-Commerce.’ *Journal of Economic Perspectives*, 15(1), pp. 69-80.

Bailey, J. and Bakos. Y., 1997, ‘An Exploratory Study of the Emerging Role of Electronic Intermediaries',’ *International Journal of Electronic Commerce* 1-3, pp. 7-20.

Baye, M. R. and Morgan. J., 2001. ‘Information Gatekeepers on the Internet and the Competitiveness of Homogeneous Product Markets,’ *American Economic Review*, 91 (3), pp. 454-474.

Baye, M. R. and Morgan. J., 2004, ‘Price Dispersion in the Lab and on the Internet, Theory and Evidence,’ *Rand Journal of Economics*, 35(3), pp. 448-466.

Baye, M. R.; Morgan, J., and Scholten, P., 2003, ‘The Value of Information in an Online Consumer Market,’ *Journal of Public Policy and Marketing*, 22 (Spring), pp. 17-25.

Baye, M. R.; Morgan, J., and Scholten, P., 2004, ‘Persistent Price Dispersion in Online Markets,’ In, D. Jansen (eds.), *The New Economy*. University of Chicago Press.

Borenstein, S., 1989, ‘Hubs and High Fares, Dominance and Market Power in the U.S. Airline Industry,’ *Rand Journal of Economics*, 20 (3), pp. 344-365.

Borenstein, S. and Rose, N., 1994, ‘Competition and Price Dispersion in the U.S. Airline Industry,’ *Journal of Political Economy*, 102(4), pp. 653-683.

Brown, J. R. and Goolsbee, A., 2002, ‘Does the Internet Make Markets More Competitive? Evidence from the Life Insurance Industry,’ *Journal of Political Economy*, 110 (3), pp. 481-507.

Brueckner, J. K., and Spiller, P. T., 1991, ‘Competition and Mergers in Airline Networks,’ *International Journal of Industrial Organization*, 9, pp. 323-342.

Brueckner, J. K., and Spiller, P. T., 1994, ‘Economies of Traffic Density in the Deregulated Airline Industry,’ *Journal of Law and Economics*, 37, pp. 379-415.

Brynjolfsson, E. and Smith, M., 1999, ‘The Great Equalizer? Consumer Choice Behavior at Internet Shopbots,’ *MIT Sloan School of Management Working Pape.*

Brynjolfsson, E. and Smith, M., 2000, ‘Frictionless Commerce? A Comparison of Internet and Conventional Retailers,’ *Management Science,* 46 (4), pp. 563-585.

Burdett, K. and Judd, K., 1983, ‘Equilibrium Price Dispersion,’ *Econometrica*, 51 (July), pp. 955-969.

Carlson, J. A. and Pescatrice, D. R., 1980, ‘Persistent Price Distributions,’ *Journal of Economics and Business,* 33(1), pp. 21-27.

Chen, J., 2006, ‘Differences in Average Prices on the Internet, Evidence from the Online Market for Air Travel,’ *Economic Inquiry,* 44(4), pp. 656-670.

Clemons, E; Hann, I. and Hitt, L., 2002, ‘Price Dispersion and Differentiation in On-Line Travel, An Empirical Investigation,’ *Management Science*, 48(4), pp. 534-549.

Diamond, P. A., 1971, ‘A Model of Price Adjustment,’ *Journal of Economic Theory*, 3, pp. 156-168.

Hopkins, E., 2006, ‘Price Dispersion’, in *New Palgrave Dictionary of Economics*, ed 2.

Lal, R. and Sarvary, M., 1999, ‘When and How is the Internet Likely to Decrease Price Competition,’ *Marketing Science*, 18(4), pp. 485-503.

Lee, D. and Prado, M., 2005, ‘The Impact of Passenger Mix on Reported ‘Hub Premiums’ in the US Airline Industry,’ *Southern Economic Journal*, 72 (2), pp. 372-394.

Lee, Z. and Gosain, S., 2002, ‘A Longitudinal Price Comparison for Music CDs in Electronics and Brick-and-Mortar Markets, Pricing Strategies in Emergent Electronic Commerce,’ *Journal of Business Strategies,* 19 (1), pp. 55-71.

Morton, F. S.; Zettelmeyer, F. and Silva-Risso, J., 2001, ‘Internet Car Retailing,’ *Journal of Industrial Economics,* 49 (4), pp. 501-519.

Morton, F. S.; Zettelmeyer, F. and Silva-Risso, J., 2003, ‘Consumer Information and Discrimination, Does the Internet Affect the Pricing of New Cars to Women and Minorities?’ *Quantitative Marketing and Economics,* 1(1), pp.65-92.

Mueller, D. C., 1986, ‘Profits in the Long Run,’ *Cambridge University Press*.

Pan, X.; Ratchford, B. T. and Shankar, V., 2002, 'Can Price Dispersion in Online Markets be Explained by Differences in E-tailer Service Quality?' *Journal of the Academy of Marketing Science,* 30(4), pp. 443-456.

Pan, X.; Ratchford, B. T. and Shankar, V., 2003, 'A Model of Retail Competition in Service and Price, Pure Play Internet vs. Bricks-and-Mortar Retailer.' *Working Paper, Indiana University, Bloomington, IN 47405.*

Pratt, J.; Wise, D. and Zeckhauser, R., 1979, 'Price Differences in Almost Competitive Markets,' *Quarterly Journal of Economics*, 93(May), pp. 189-211.

Ravenscraft, D., 1983, 'Structure-profit Relationships at the Line of Business and Industry Level,' *Review of Economics and Statistics*, 65, pp. 22-31.

Reinganum, J. F., 1979, 'A Simple Model of Equilibrium Price Dispersion,' *Journal of Political Economy*, 87, pp. 851-858.

Rothschild, M., 1974, 'Searching for the Lowest Price When the Distribution of Prices Is Unknown,' *Journal of Political Economy*, 82(4), pp. 689-711.

Salop, S. and Stiglitz, J., 1977, 'Bargains and Rip-offs, A Model of Monopolistically Competitive Price Dispersion,' *The Review of Economic Studies,* 44 (3), pp. 493-510.

Scholten, P. A. and Smith, A., 2002, 'Price Dispersion Then and Now, Evidence from Retail and E-Tail Markets,' *Advances in Microeconomics, Economics of the Internet and e-Commerce*, Vol. 11, pp. 63-88.

Shilony, Y., 1977, 'Mixed Pricing in Oligopoly,' *Journal of Economic Theory*, 14, pp. 373-388.

Smith, B.; Rao, B. and Ratliff, R., 2001, 'E-Commerce and Operations Research in Airline Planning, Marketing, and Distribution,' *Interfaces,* 31(2), pp. 37-55.

Smith, M.; Bailey, J. and Brynjolfsson, E., 1999. 'Understanding Digital Markets, Review and Assessment,' in *Understanding the Digital Economy,* Brynjolfsson and Kahin eds., MIT Press, Cambridge.

Sorensen, A., 2000, 'Equilibrium Price Dispersion in Retail Markets for Prescription Drugs,' *Journal of Political Economy,* 108(4), pp. 833-850.

Stahl, D., 1989, 'Oligopolistic Pricing with Sequential Consumer Search,' *American Economic Review*, 79, September 1989, pp. 700-712.

Stavins, J., 2001, ‘Price Discrimination in the Airline Market, The Effect of Market Concentration.’ *Review of Economics and Statistics*. 83 (1), pp. 200-202.

Stigler, J.G., 1961, ‘The Economics of Information,’ *Journal of Political Economy*, 69(3), pp. 213-225.

Varian, H. R., 1980, ‘A Model of Sales,’ *American Economic Review*, 70 (4), 651-659.

Verlinda, J. A. and Lane, L., 2004, ‘The Effect of the Internet on Pricing in the Airline Industry,’ *Social Science Research Network*, abstract #965788.

Waldeck, R., 2002, ‘Pricing and Price Dispersion in E-commerce’, manuscript prepared for *International Conference on Electronic Commerce Research, ICECR-5*. (Website: http://www-eco.enst-bretagne.fr/Recherche/Biblio/Roger/pricing%20in%20e-commerce.pdf)

Wilde, L. and Schwartz, A., 1979, ‘Equilibrium Comparison Shopping,’ *Review of Economic Studies*, 46, pp. 543-553.

APPENDIX A

I. Data description, construction of variables and expected effects

We sketch below a detailed description of the variables used and how they were constructed. The final data set used for the analysis has been comprised from three different data sets. The first data set includes contemporaneous online and offline transaction data from the fourth quarter of 2004. However, our period includes some of the peak travel period, particularly Thanksgiving, Christmas and New Years. To sidestep the problems of pricing during these peak travel periods, we dropped transactions for travel during the Thanksgiving week. We also kept transaction which included departure and return within the 22nd of December, 2004. Thus we do not include itineraries involving travel during the last week of the year, since pricing can be different for these periods.

This transaction data comes from one of the major computer reservation systems. Unfortunately, due to confidentiality reasons, they did not provide us with the ticket restrictions. To overcome this limitation, we collected computer reservation system data by gathering the same from one of the local travel agents. The travel agents systems can access historical data for a year. However, due to the time difference between the actual period for which we had data and the data that we could collect, we could obtain a subset of the prices and their characteristics that were offered for the last quarter of 2004, since much of the data was taken out from the reservation systems in a random manner. We matched our transaction data to the travel agents data to obtain the restrictions on the individual tickets. To overcome, the data limitation problem arising from the sub-set of the data that we could collect, we adopted a matching rule. If the two prices from the data sets matched within a 2 percent range, we assigned it as a match. We are thereby assuming, that for a ticket priced at $150 will be qualitative similar to one priced at $147 or $153. We however, took full precaution that the other matching criteria like carrier, booking class and coach class, the day of the week of travel (in case we matched it with a ticket that has a travel day restriction) and the advance purchase requirement were matched in both the data sets.

Following Borenstein (1989) and Borenstein and Rose (1994), we include itineraries which has at most of one break (stop-over) in either direction. The prices are for roundtrip fares. For the one-way itineraries the fares are multiplies by two. We exclude all itineraries which are open-jaw and circular trip tickets. This study includes tickets which are operated by American Airlines, Continental, Delta, Northwest, US Airways, United Airlines, Frontier, Air Tran, Spirit, Alaska, American Mid-west, Sun Country, Frontier Airlines and American Trans Air.

The following variables are included:

Refund: equals 1 if the ticket is refundable else 0 (sign: positive)

Advance: the advance purchase requirement required on a ticket. This is usually 1, 3, 5, 7, 10, 14, 21 and 30 (sign: negative)

Days prior to departure: the number of days before departure the ticket was bought. Given our data, we noticed some really low fares very close to the departure date; even on the day of departure while some really high fares more than a month in advance. So the relationship between days prior to departure and price paid is indeterminate. (sign: ?)

Online: equal to 1 if the ticket was bought online and 0 if purchased offline (sign: negative)

Direct: equal to 1 if the itinerary did not involve a change of plane or stop-over and 0 otherwise. This is an indicator for non-stop itinerary (sign: positive)

Roundtrip: equal to 1 if the itinerary was for a roundtrip travel and 0 otherwise (sign: negative)

Saturday nigh stay-over: equal to 1 if the itinerary involved a Saturday stay over and 0 otherwise. This was created by using the departure and the return date indicated in the transaction data (sign: negative)

First class: equal to 1 if any segment of the itinerary involved a travel in the first class coach and 0 otherwise (sign: positive)

Business class: equal to 1 if any segment of the itinerary involved a travel in full coach fare class or business class (not first class) and 0 otherwise (sign: positive)

Travel Restriction: equal to 1 if the ticket required a travel day restriction and 0 otherwise. This primarily requires, that the ticket bought (price) is valid only if the individual travels during certain days of the week, say, Tuesday or Thursday (sign: negative)

Minimum stay restriction: equal to 1 if the ticket required a minimum stay required and 0 otherwise (sign: negative)

Maximum stay restriction: equal to 1 if the ticket required a maximum stay requirement is imposed on the ticket and 0 otherwise (sign: negative)

Distance (in logs): non-stop mileage between the two endpoint airports on a route (sign: positive)

Temperature difference (in logs): the absolute difference in the average January temperature between the origin and destination of the individual itinerary (sign: negative)

Hub: equal to 1 if any endpoint airport of the route is a hub airport for the operating carrier (sign: positive)

Slots: equal to 1 if any endpoint airport has restricted slots. This includes LGA, JFK and DCA.

Average load factor at time of purchase: this is the load factor averaged over each individual segment involved in the itinerary at the time when the ticket was purchased. We had information regarding the flight numbers for travel in each segment of an individual itinerary. We used this information along with data from Official Airline Guide (OAG) to calculate the total number of seats on each of these flights scheduled for departure on a particular date. From our transaction data, we calculated the total number of seats that were sold on that flight the day before an individual transaction for that flight took place. That is, for a ticket involving a travel on flight 66 on American Airlines from DFW-ORD on October 10, 2004, and being bought on October 9th, we calculate how many seats were sold of flight 66 departing on October 10th was sold till October 8th. Since we cannot, observe the order of transaction taking place on the same day (October 9th) we assume that all tickets being bought October 9th for October 10th flight will face the same load factor as of October 8th. This is the closest approximation possible to calculate the contemporaneous load factor facing an individual ticket at the time of transaction. For all the segments involved in the itinerary, we average these load factors to calculate the average load factor associated with each transaction. We expect this variable to have a positive coefficient; since individuals will face a higher price if he travels on flights which are full (have a high demand) (sign: positive)

Departure at peak time: equals to 1 if the individual itinerary involves departure at a peak time (between 8-10am or 3-7pm). Given the flight numbers, we use information from OAG to determine the local departure time (sign: positive)

Return at peak time: equals to 1 if the individual itinerary involves return at a peak time (between 8-10am or 3-7pm). For one way tickets, this is equal to 0 (sign: positive)

Low cost route (lcr): takes a value equal to 1 if a low cost carrier (other than Southwest) operates on that route and 0 otherwise. The presence of the low cost carrier is expected to induce competition among operating airlines driving the prices down (sign: negative)

Southwest (sw): takes a value equal to 1 if Southwest airlines operate on that route and 0 otherwise. It is well accepted that presence of Southwest airlines induces significant price reductions on the route owing to cost advantage of Southwest (sign: negative)

Average population (in logs): the average population at the two endpoints of the route (Source: US Census 2003). On one hand, higher population between the endpoint airports of a route can create more demand such that price increases. On the other hand, airlines can have more flights on routes where there is more demand such that prices can be lower (sign :?)

Average per-capita income (in logs): the average per capita income at the two endpoints of the route (Source: US Census 2003) (sign: positive)

Market share: calculated as the proportion of passengers accounted by a carrier on a route. T-100 segment data is used to calculate this share. If there is not a complete umbrella effect from the market power of a dominant firm, then holding the market concentration constant, an increase in market share is expected to increase the prices (sign: positive).

Herfindahl Index (HHI): sum of the square of the market shares of each of the carriers operating on a route. To the extent that the dominant firm's high prices create an umbrella that allows a few firms in a concentrated market to collude easily, then increases in the concentration will increase the prices. If however, a dominant firm on a route has a competitive advantage owing to cost structure, advertising, marketing or other means, then it could possibly reduce the profit maximizing prices of the other firms (sign: ?)

Internet share: share of all online transactions to the total transactions on a route (sign: negative)

Internet share*Online: INT interacted with the online dummy (sign: ?)

Geoshare: given by ($\sqrt{}ENP_{x1}$. ENP_{x2})/ ($\sqrt{}ENP_{y1}$. ENP_{y2}) where y indexes all airlines, x the observed airline and ENP_{y1} and ENP_{y2} are airline y's average enplanements at the two enpoints airports during the fourth quarter of 2004.

Xtherf: is the square of the fitted value of for market share (from its first stage regression) plus the rescaled sum of the squared of all other carriers' share. This is given by:
xtherf = (predicted market share)^2 + [(HHI-market share^2)/(1-market share)^2]*(1-predcited market share)^2 . See discussion in Borenstein (1989) and/or Borenstein and Rose (1994).

Departure day of the week: indicates the departure day of the week. It takes a value of 0 to 6, where 0 represents a Sunday and 6 a Saturday. In all estimates Sunday is treated as the base group.

Return day of the week: indicates the return day of the week. It takes a value of 0 to 6, where 0 represents a Sunday and 6 a Saturday. In all estimates Sunday is treated as the base group. Also, for one way tickets, this takes a value of 0.

The rest of this appendix reports some of the additional results that were referred to in the main text in Chapter III but were not reported. Table 1A reports the estimation results of the effects of the internet on the ticket prices, when we include the route effects and not the market structure variables, namely, market share of the carrier on a route, Herfindahl index and others. Table A2 includes the complete list of all the 150 city-pair routes that has been used in the analysis in Chapter III. Finally, Table A3 reports results of the sensitivity analysis as discussed in section 7 of Chapter III. The results in Table A3 clearly suggest that the direct and indirect effects of the internet on average prices are robust even when we include the five percent matched samples. This clearly demonstrates that the effects that we found in the main analysis was not the result of a sample selection bias but holds good in general.

Table A1
The Effects of Online When Route Fixed Effects Are Included

	lfare
Non-refundable	-0.4383
	(269.67)**
Days prior to departure ticket purchased	0.0004
	(14.67)**
Advance Purchase Requirement:	
One day advance	-0.2426
	(66.63)**
Three day advance	-0.0449
	(27.39)**
Five day advance	-0.5389
	(38.27)**
Seven day advance	-0.1705
	(121.60)**
Ten day advance	-0.2067
	(67.22)**
Fourteen day advance	-0.2593
	(162.75)**
Twenty-one day advance	-0.3016
	(91.99)**
Thirty day advance	-0.1294
	(16.56)**
Direct	0.0545
	(13.28)**
Online	-0.1143
	(82.13)**
Roundtrip	-0.1064
	(48.73)**
Saturday stay-over	-0.1166
	(62.23)**
First class	0.6987
	(362.47)**
Business class	0.3156
	(192.71)**
Travel restriction	-0.2749
	(264.53)**
Minimum stay requirement	-0.0084
	(5.65)**
Maximum stay requirement	-0.0466
	(30.86)**
HUB	0.0225
	(8.49)**
Average load factor at purchase	0.2602
	(39.82)**

Table A1, Continued

	Log(fare)
Peak time of Day:	
Departure at peak time	0.008
	(8.74)**
Return at peak time	0.017
	(15.33)**
Airline Fixed Effects (American Airlines omitted):	
Continental	0.0009
	(0.32)
Delta	-0.0006
	(0.27)
Northwest	-0.0139
	(4.02)**
United Airways	0.098
	(64.64)**
US Air	-0.0318
	(11.05)**
Frontier	-0.2295
	(57.08)**
Alaska	-0.0433
	(4.64)**
Hawaiian Airlines	0.0035
	(0.23)
America Mid-west	-0.2206
	(54.74)**
American Trans Air	-0.2106
	(14.70)**
Midwest Express	-0.0396
	(2.64)**
Air Tran	-0.3693
	(34.39)**
Spirit	-0.2157
	(36.12)**
Sun County	-0.4723
	(55.54)**
Departure Day of Week (Sunday omitted):	
Monday	0.0011
	(0.71)
Tuesday	0.0012
	(0.73)
Wednesday	0.0023
	(1.34)
Thursday	0.0085
	(4.67)**
Friday	0.0149
	(7.61)**
Saturday	-0.0642
	(30.25)**

Table A1, Continued

	Log(fare)
Return day of week (Sunday omitted):	
Monday	-0.032
	(15.08)**
Tuesday	-0.0258
	(11.55)**
Wednesday	-0.0338
	(15.02)**
Thursday	-0.0254
	(11.25)**
Friday	-0.0345
	(15.47)**
Saturday	-0.0638
	(23.60)**
Constant	6.1247
	(854.84)**
Route Effects	Yes
Observations	523618
R-squared	0.8

Absolute value of t statistics in parentheses
* significant at 5%; ** significant at 1%

Table A2
List of City Pairs Used

Routes	Routes
Atlanta (ATL)-Boston (BOS)	Chicago (ORD) – Orange County (SNA)
Atlanta (ATL)-Cincinnati (CVG)	Chicago (MDW) – Detroit (DTW)
Atlanta (ATL)- Fort Lauderdale (FLL)	Cleveland (CLE) – Chicago (MDW)
Atlanta (ATL)-Dulles, DC (IAD)	Cleveland (CLE)– Chicago O' Hare (ORD)
Atlanta (ATL)-Houston (IAH)	Cincinnati (CVG)–O' Hare (ORD)
Atlanta (ATL)-Los Angeles (LAX)	Columbus (CMH) – La Guardia (LGA)
Atlanta (ATL)-La Guardia (LGA)	Dallas (DFW) – Atlanta (ATL)
Atlanta (ATL)- Orlando (MCO)	Dallas (DFW) – Denver (DEN)
Atlanta (ATL)- Memphis (MEM)	Dallas (DFW) – Washington (IAD)
Atlanta (ATL) – Miami (MIA)	Dallas (DFW)- Houston (IAH)
Atlanta (ATL)-New Orleans (MSY)	Dallas (DFW) – Los Angeles (LAX)
Atlanta (ATL) – Chicago O' Hare (ORD)	Dallas (DFW) – Long Beach (LGB)
Atlanta (ATL)- Philadelphia (PHL)	Dallas (DFW) – Kansas City (MCI)
Atlanta (ATL)-Tampa (TPA)	Dallas (DFW) – Chicago (ORD)
Baltimore (BWI) – Atlanta (ATL)	Dallas (DFW) – Phoenix (PHX)
Baltimore (BWI) – Cleveland (CLE)	Denver (DEN) – Atlanta (ATL)
Baltimore (BWI) – Dallas (DFW)	Denver (DEN) – Boston (BOS)
Baltimore (BWI)- Fort Lauderdale (FLL)	Denver (DEN) – Washington (DCA)
Baltimore (BWI) – Los Angeles (LAX)	Denver (DEN) – Newark (EWR)
Baltimore (BWI)- Orlando (MCO)	Denver (DEN) – Houston (IAH)
Boston (BOS) – Baltimore (BWI)	Denver (DEN) – New York (LGA)
Boston (BOS)- Charlotte (CLT)	Denver (DEN) – Kansas City (MCI)
Boston (BOS)- Washington (DCA)	Denver (DEN) – Orlando (MCO)
Boston (BOS) – Dallas (DFW)	Denver (DEN) – Portland (PDX)
Boston (BOS) – Detroit (DTW)	Denver (DEN) – Philadelphia (PHL)
Boston (BOS) – Los Angeles (LAX)	Denver (DEN) – Phoenix (PHX)
Boston (BOS) – Philadelphia (PHL)	Denver (DEN) – St. Louis (STL)
Boston (BOS) – Pittsburgh (PIT)	Denver (DEN) – Tampa (TPA)
Boston (BOS) – Fort Myers (RSW)	Detroit (DTW) – Atlanta (ATL)
Boston (BOS) – Tampa (TPA)	Detroit (DTW) – Baltimore (BWI)
Charlotte (CLT) – Orlando (MCO)	Detroit (DTW) – Dallas (DFW)
Chicago (ORD) – Boston (BOS)	Detroit (DTW) – Newark (EWR)
Chicago (ORD) – Baltimore (BWI)	Detroit (DTW) – Fort Lauderdale (FLL)
Chicago (ORD) – Charlotte (CLT)	Detroit (DTW) – Las Vegas (LAS)
Chicago (ORD) – Denver (DEN)	Detroit (DTW) – Orlando (MCO)
Chicago (ORD) – Washington (IAD)	Detroit (DTW) – Chicago (ORD)
Chicago (ORD)- New York (LGA)	Fort Lauderdale (FLL) – Boston (BOS)
Chicago (ORD) – Miami (MIA)	Fort Lauderdale (FLL)- Chicago (ORD)
Chicago (ORD) – Minneapolis (MSP)	Hartford (BDL) – Washington (DCA)
Chicago (ORD) – New Orleans (MSY)	Hartford (BDL) – Chicago O' Hare (ORD)
Chicago (ORD) – Omaha (OMA)	Honolulu (HNL) – Los Angeles (LAX)
Chicago (ORD) – Ft. Myers (RSW)	Houston (IAH) – New Orleans (MSY)
Chicago (ORD) – San Diego (SAN)	Houston (IAH) – Chicago (ORD)
Las Vegas (LAS) – Burbank (BUR)	New York (LGA) – Cincinnati (CVG)
Las Vegas (LAS) – Los Angeles (LAX)	New York (LGA)- Dallas (DFW)
Las Vegas (LAS) – Chicago (ORD)	New York (LGA) – Detroit (DTW)
Long Beach (LGB) – Dallas (DFW)	New York (LGA)- Houston (IAH)

Table A2, Continued

Routes	Routes
Los Angeles (LAX) – Denver (DEN)	New York (LGA) – Palm Beach, FL (PBI)
Los Angeles (LAX) – Detroit (DTW)	Oakland (OAK) – Denver (DEN)
Los Angeles (LAX) – Houston (IAH)	Oakland (OAK) – Seattle (SEA)
Los Angeles (LAX)- Miami (MIA)	Ontario (ONT) – Denver (DEN)
Los Angeles (LAX)- Chicago (ORD)	Orlando (MCO) – Washington (DCA)
Los Angeles (LAX) – Philadelphia (PHL)	Orlando (MCO) – Dallas (DFW)
Los Angeles (LAX) – Reno (RNO)	Orlando (MCO)- New York (LGA)
Los Angeles (LAX) – Tampa (TPA)	Palm Beach (PBI) – Boston (BOS)
Miami (MIA) – New York (LGA)	Philadelphia (PHL) – Chicago (ORD)
Miami (MIA) – Boston (BOS)	Philadelphia (PHL) – Palm Beach (PBI)
Miami (MIA)- Newark (EWR)	Phoenix (PHX) – Minneapolis (MSP)
Milwaukee (MKE) – Minneapolis (MSP)	Phoenix (PHX) – Ontario (ONT)
Minneapolis (MSP) – Denver (DEN)	Pittsburgh (PIT) – New York (LGA)
Minneapolis (MSP) – Dallas (DFW)	Pittsburgh (PIT) – Chicago (ORD)
Minneapolis (MSP) – Detroit (DTW)	Portland (PDX) – Las Vegas (LAX)
Minneapolis (MSP) – Los Angeles (LAX)	Portland (PDX) – Los Angeles (LAX)
Minneapolis (MSP) – New York (LGA)	Portland (PDX) – Oakland (OAK)
Minneapolis (MSP) – Chicago (MDW)	St. Louis (STL) – Los Angeles (LAX)
Newark (EWR) – Minneapolis (MSP)	Sacramento (SMF) – Los Angeles (LAX)
Newark (EWR) – Chicago (ORD)	Salt Lake City (SLC) – Denver (DEN)
Newark (EWR) – Atlanta (ATL)	San Francisco (SFO) – Boston (BOS)
Newark (EWR) – Boston (BOS)	San Francisco (SFO) – Dallas (DFW)
Newark (EWR) – Los Angeles (LAX)	San Jose (SJC) – Denver (DEN)
New Orleans (MSY) – New York (LGA)	Tampa (TPA) – New York (LGA)
New York (JFK) – Los Angeles (LAX)	Washington (DCA) – Atlanta (ATL)
New York (LGA) – Boston (BOS)	Washington (DCA) – Dallas (DFW)
New York (LGA) – Cleveland (CLE)	Washington (DCA)-La Guardia (LGA)
New York (LGA) – Charlotte (CLT)	Washington (DCA)- Chicago (ORD)

Table A3
Internet Purchase, Internet Share and Potential Savings from Internet Purchase in High-Low Internet Usage Markets for Observations Matched Within Five Percent Range

	(1) Log(Fare) (OLS)	(2) Log(Fare) (IV)	(3) Log(Fare) (OLS)	(4) Log(Fare) (IV)	(5) Log(Fare) (OLS)	(6) Log(Fare) (IV)
Market Structure Variables:						
Market share	0.0786	0.2948	0.0847	0.2985	0.0833	0.297
	(26.31)**	(32.71)**	(28.49)**	(33.30)**	(27.99)**	(33.08)**
HHI	0.0391	-0.263	0.0102	-0.2613	0.0118	-0.2592
	(10.60)**	(27.05)**	(2.78)**	(27.03)**	(3.20)**	(26.76)**
HUB	0.1438	0.1249	0.1333	0.1123	0.1332	0.1122
	(110.14)**	(76.63)**	(102.18)**	(68.85)**	(102.13)**	(68.84)**
Slot constrained airport	0.1192	0.1205	0.1171	0.1149	0.1169	0.1147
	(83.86)**	(80.00)**	(82.83)**	(76.76)**	(82.68)**	(76.65)**
Internet Variables:						
Online	-0.1537	-0.149	-0.146	-0.1409	-0.1796	-0.1666
	(114.36)**	(109.33)**	(108.93)**	(103.65)**	(60.11)**	(54.81)**
Internet share			-0.6039	-0.6357	-0.6326	-0.6574
			(90.75)**	(94.12)**	(89.96)**	(92.48)**
Internet share*Online					0.1449	0.1108
					(12.61)**	(9.54)**
Advance Purchase Requirement: **(No advance purchase required omitted)**						
One day advance	-0.4397	-0.4378	-0.4065	-0.4046	-0.4059	-0.4041
	(133.88)**	(132.49)**	(123.65)**	(122.33)**	(123.45)**	(122.19)**
Three day advance	-0.005	0.0001	0.0019	0.0064	0.0023	0.0068
	(3.11)**	(0.06)	(1.16)	(4.00)**	(1.47)	(4.22)**
Five day advance	-0.6315	-0.6276	-0.5879	-0.5854	-0.5888	-0.5861
	(40.63)**	(40.19)**	(38.00)**	(37.68)**	(38.06)**	(37.73)**

Table A3, Continued

	(1) Log(Fare) (OLS)	(2) Log(Fare) (IV)	(3) Log(Fare) (OLS)	(4) Log(Fare) (IV)	(5) Log(Fare) (OLS)	(6) Log(Fare) (IV)
Seven day advance	-0.1729	-0.1724	-0.1701	-0.1697	-0.1697	-0.1695
	(127.77)**	(126.77)**	(126.30)**	(125.53)**	(126.04)**	(125.32)**
Ten day advance	-0.2375	-0.2309	-0.2393	-0.2349	-0.2385	-0.2343
	(81.40)**	(78.65)**	(82.46)**	(80.47)**	(82.17)**	(80.26)**
Fourteen day advance	-0.2372	-0.2323	-0.241	-0.2375	-0.2405	-0.2372
	(155.44)**	(151.10)**	(158.72)**	(155.39)**	(158.34)**	(155.12)**
Twenty-one day advance	-0.2591	-0.258	-0.2571	-0.2568	-0.2564	-0.2563
	(80.91)**	(80.20)**	(80.72)**	(80.28)**	(80.48)**	(80.10)**
Thirty day advance	0.0293	0.0317	0.0221	0.0239	0.0238	0.0252
	(4.10)**	(4.41)**	(3.11)**	(3.35)**	(3.35)**	(3.53)**
Other Ticket Characteristics:						
Non-refundable ticket	-0.4026	-0.4082	-0.4019	-0.4059	-0.4016	-0.4057
	(268.23)**	(269.48)**	(269.11)**	(269.62)**	(268.91)**	(269.43)**
Days prior to departure ticket purchased	-0.0004	-0.0005	-0.0002	-0.0003	-0.0002	-0.0003
	(16.57)**	(18.60)**	(8.69)**	(10.12)**	(9.50)**	(10.73)**
Saturday stay-over	-0.1185	-0.1189	-0.1146	-0.1148	-0.1151	-0.1152
	(64.98)**	(64.87)**	(63.10)**	(62.98)**	(63.39)**	(63.21)**
Travel restriction requirement	-0.3052	-0.3073	-0.3011	-0.3019	-0.3011	-0.3019
	(314.29)**	(313.35)**	(311.27)**	(309.44)**	(311.29)**	(309.46)**
Minimum stay requirement	0.0132	0.0171	0.0102	0.0135	0.0102	0.0134
	(10.00)**	(12.79)**	(7.73)**	(10.16)**	(7.71)**	(10.13)**
Maximum stay requirement	0.0271	0.0312	0.0304	0.0339	0.0306	0.0341
	(19.55)**	(22.27)**	(21.97)**	(24.34)**	(22.15)**	(24.48)**
First class	0.7155	0.7137	0.7042	0.7016	0.7035	0.7011
	(377.53)**	(374.45)**	(372.65)**	(369.31)**	(372.20)**	(369.01)**
Business class	0.3691	0.3778	0.3758	0.3825	0.3757	0.3823
	(248.25)**	(249.99)**	(253.73)**	(254.11)**	(253.67)**	(254.02)**
Remaining Ticket Characteristics:						
Direct flight	0.0237	-0.0165	0.0203	-0.0181	0.0206	-0.0177

Table A3, Continued

	(1) Log(Fare) (OLS)	(2) Log(Fare) (IV)	(3) Log(Fare) (OLS)	(4) Log(Fare) (IV)	(5) Log(Fare) (OLS)	(6) Log(Fare) (IV)
	(6.21)**	(4.03)**	(5.33)**	(4.44)**	(5.43)**	(4.35)**
Roundtrip ticket	-0.0884	-0.0904	-0.0933	-0.0961	-0.0926	-0.0956
	(41.34)**	(41.97)**	(43.81)**	(44.83)**	(43.50)**	(44.58)**
Flight load factor at purchase	0.2016	0.1779	0.1761	0.1517	0.1726	0.1491
	(30.60)**	(26.64)**	(26.84)**	(22.82)**	(26.29)**	(22.42)**
<u>Peak Time of Day:</u>						
Departure at peak time	0.0172	0.0186	0.0173	0.0186	0.0173	0.0186
	(17.72)**	(19.11)**	(17.90)**	(19.24)**	(17.91)**	(19.25)**
Return at peak time	0.0279	0.0286	0.0272	0.0279	0.0272	0.0279
	(24.97)**	(25.48)**	(24.44)**	(24.96)**	(24.40)**	(24.93)**
<u>Other Route Specific Characteristics:</u>						
Low cost carrier on route	-0.1148	-0.1166	-0.0897	-0.0904	-0.0892	-0.09
	(116.49)**	(117.06)**	(88.00)**	(87.94)**	(87.40)**	(87.45)**
Southwest Airlines	-0.1977	-0.1894	-0.1691	-0.1606	-0.1701	-0.1614
	(99.45)**	(94.25)**	(84.42)**	(79.18)**	(84.87)**	(79.48)**
<u>Other Route Level Variables:</u>						
Distance (log)	0.3839	0.3773	0.4223	0.4166	0.423	0.4171
	(451.46)**	(395.58)**	(446.43)**	(403.87)**	(446.46)**	(403.04)**
Absolute Temperature	-0.0068	-0.0075	-0.0052	-0.0056	-0.0052	-0.0057
Difference (Log)	(18.12)**	(19.84)**	(13.92)**	(15.06)**	(14.02)**	(15.13)**
Average population (Log)	-0.026	-0.0271	-0.0376	-0.0377	-0.0375	-0.0376
	(38.52)**	38.03)**	(54.94)**	(52.22)**	(54.84)**	(52.12)**
Average per capita	0.2036	0.206	0.096	0.0944	0.0943	0.0932
Income (Log)	(36.44)**	(36.65)**	(16.89)**	(16.52)**	(16.59)**	(16.30)**
<u>Departure Day of Week (Sunday omitted):</u>						
Monday	0.0027	0.0029	0.0024	0.0026	0.0023	0.0025

Table A3, Continued

	(1) Log(Fare) (OLS)	(2) Log(Fare) (IV)	(3) Log(Fare) (OLS)	(4) Log(Fare) (IV)	(5) Log(Fare) (OLS)	(6) Log(Fare) (IV)
	(1.64)	(1.78)	(1.46)	(1.61)	(1.4)	(1.57)
Tuesday	0.0051	0.0056	0.0043	0.0047	0.0042	0.0047
	(3.02)**	(3.32)**	(2.54)*	(2.80)**	(2.50)*	(2.77)**
Wednesday	0.0047	0.0057	0.0046	0.0055	0.0044	0.0053
	(2.65)**	(3.20)**	(2.64)**	(3.12)**	(2.52)*	(3.02)**
Thursday	0.0115	0.0128	0.0116	0.0128	0.0115	0.0127
	(6.21)**	(6.91)**	(6.33)**	(6.95)**	(6.24)**	(6.89)**
Friday	0.0146	0.0151	0.0143	0.0146	0.0145	0.0148
	(7.36)**	(7.58)**	(7.25)**	(7.38)**	(7.36)**	(7.46)**
Saturday	-0.0675	-0.0673	-0.0665	-0.0665	-0.0665	-0.0665
	(31.94)**	(31.69)**	(31.63)**	(31.49)**	(31.62)**	(31.49)**
Return Day of the Week (Sunday omitted):						
Monday	-0.048	-0.0482	-0.0466	-0.0466	-0.0474	-0.0472
	(23.60)**	(23.58)**	(23.01)**	(22.92)**	(23.43)**	(23.24)**
Tuesday	-0.0379	-0.0379	-0.0366	-0.0363	-0.0378	-0.0372
	(17.53)**	(17.42)**	(17.00)**	(16.82)**	(17.54)**	(17.23)**
Wednesday	-0.0464	-0.0465	-0.0445	-0.0443	-0.0456	-0.0452
	(21.27)**	(21.22)**	(20.48)**	(20.32)**	(20.98)**	(20.70)**
Thursday	-0.0342	-0.0349	-0.0323	-0.0326	-0.0333	-0.0334
	(15.65)**	(15.88)**	(14.82)**	(14.93)**	(15.30)**	(15.29)**
Friday	-0.0408	-0.0416	-0.0384	-0.0389	-0.0394	-0.0397
	(18.85)**	(19.16)**	(17.86)**	(18.04)**	(18.32)**	(18.38)**
Saturday	-0.0752	-0.076	-0.0727	-0.0733	-0.0736	-0.074
	(28.97)**	(29.17)**	(28.16)**	(28.28)**	(28.52)**	(28.55)**
Carrier Fixed Effects (American Airlines omitted):						
Continental	-0.045	-0.0415	-0.053	-0.0502	-0.0534	-0.0505
	(26.71)**	(24.44)**	(31.57)**	(29.68)**	(31.81)**	(29.87)**
Delta	-0.1116	-0.123	-0.1169	-0.1282	-0.1168	-0.1281

Table A3, Continued

	(1) Log(Fare) (OLS)	(2) Log(Fare) (IV)	(3) Log(Fare) (OLS)	(4) Log(Fare) (IV)	(5) Log(Fare) (OLS)	(6) Log(Fare) (IV)
	(76.28)**	(80.50)**	(80.26)**	(84.23)**	(80.21)**	(84.16)**
Northwest	0.0742	0.0622	0.1214	0.1114	0.1215	0.1115
	(37.71)**	(30.32)**	(59.97)**	(53.20)**	(60.02)**	(53.23)**
United Airways	0.1076	0.0989	0.1345	0.1293	0.135	0.1297
	(81.24)**	(73.32)**	(99.55)**	(94.58)**	(99.90)**	94.84)**
US Air	-0.1123	-0.132	-0.0929	-0.1127	-0.0934	-0.1131
	(61.15)**	(63.66)**	(50.51)**	(54.57)**	(50.77)**	(54.78)**
Frontier Airlines	0.0694	0.0835	0.1241	0.143	0.1213	0.1409
	(22.23)**	(25.79)**	(39.25)**	(43.49)**	(38.27)**	(42.66)**
Alaska	-0.0986	-0.1353	-0.0379	-0.0677	-0.0372	-0.0671
	(21.90)**	(29.07)**	(8.37)**	(14.54)**	(8.22)**	(14.40)**
Hawaiian Airlines	0.3265	0.2879	0.3027	0.2705	0.3033	0.2711
	(22.06)**	(19.33)**	(20.55)**	(18.25)**	(20.59)**	(18.30)**
America West	-0.0592	-0.0508	-0.0234	-0.0079	-0.0229	-0.0075
	(16.82)**	(13.72)**	(6.65)**	(2.14)*	(6.50)**	(2.02)*
American Trans Air	0.029	0.0138	0.0348	0.0189	0.037	0.0206
	(2.56)*	(1.22)	(3.10)**	(1.67)	(3.29)**	(1.82)
Midwest Express Airlines	0.5008	0.6671	0.484	0.6427	0.482	0.6409
	(36.07)**	(43.91)**	(35.03)**	(42.58)**	(34.89)**	42.43)**
Air Tran	-0.3341	-0.279	-0.3925	-0.3428	-0.3937	-0.3439
	41.96)**	(33.86)**	(49.38)**	(41.80)**	(49.54)**	(41.92)**
Spirit	-0.2288	-0.2044	-0.1716	-0.1433	-0.1768	-0.1473
	(47.83)**	(41.46)**	(35.74)**	(28.92)**	(36.70)**	29.57)**
Sun County	-0.1108	-0.0324	-0.0588	0.0205	-0.0649	0.0156
	16.31)**	(4.35)**	(8.66)**	(2.75)**	(9.54)**	(2.09)*
Constant	1.9238	2.0661	3.0677	3.2206	3.0851	3.2328
	(31.43)**	(33.50)**	(49.32)**	(51.34)**	(49.60)**	(51.53)**
Observations	794134	794134	794134	794134	794134	794134
R-squared	0.690	0.680	0.690	0.690	0.690	0.690

APPENDIX B

Proposition 1: *The expected price decreases as the number of high intensity searchers in each of the individual markets rises.*

Proof:

$$E(P^0) = P_0 + \int_{P_0}^{r^0}\left[\frac{(1-\mu_0)(r^0-P^0)}{nP^0\mu_0}\right]^{\frac{1}{n-1}} dP^0$$

$$\text{or, } \frac{\partial E(P^0)}{\partial \mu_0} = \frac{\partial P^0}{\partial \mu_0} + \frac{\partial I}{\partial \mu_0}$$

$$\text{where } I = \int_{P_0}^{r^0}\left[\frac{(1-\mu_0)(r^0-P^0)}{nP^0\mu_0}\right]^{\frac{1}{n-1}} dP^0. \text{ Now, } I = \int_{P_0}^{r^0}\left[\frac{(1-\mu_0)(r^0-P^0)}{nP^0\mu_0}\right]^{\frac{1}{n-1}} dP^0$$

$$\text{or, } I = \left[\frac{1-\mu_0}{n\mu_0}\right]^{1/n-1}\int_{P_0}^{r^0}\left[\frac{(1-\mu_0)(r^0-P^0)}{nP^0\mu_0}\right]^{\frac{1}{n-1}} dP^0 = u.v$$

$$\text{where } u = \left[\frac{1-\mu_0}{n\mu_0}\right]^{1/n-1} \text{ and } v = \int_{P_0}^{r^0}\left[\frac{(1-\mu_0)(r^0-P^0)}{nP^0\mu_0}\right]^{\frac{1}{n-1}} dP^0$$

$$\text{Now, } u = \left[\frac{1-\mu_0}{n\mu_0}\right]^{1/n-1}$$

Therefore,

$$\frac{\partial u}{\partial \mu_0} = \left[\frac{1}{n}\right]^{1/n-1}\left[-(1/n-1)(\mu_0)^{\frac{-1}{n-1}-1}(1-\mu_0)^{\frac{1}{n-1}} + (1/n-1)(\mu_0)^{\frac{-1}{n-1}}(1-\mu_0)^{\frac{1}{n-1}-1}\right]$$

$$= (1/n)^{\frac{1}{n-1}}\left[-\left(\frac{1}{n-1}\right)\left(\frac{(1-\mu_0)^{\frac{1}{n-1}}}{(\mu_0)^{\frac{1}{n-1}+1}} + (1-\mu_0/\mu_0)^{\frac{1}{n-1}}\frac{1}{1-\mu_0}\right)\right] \Rightarrow \frac{\partial u}{\partial \mu_0} < 0$$

$$\text{Now, } v = \int_{P_0}^{r^0}\left[\frac{(1-\mu_0)(r^0-P^0)}{nP^0\mu_0}\right]^{\frac{1}{n-1}} dP^0,$$

Hence,

$$\frac{\partial v}{\partial \mu_0} = \frac{\partial}{\partial \lambda_0}\left[\int_{P_0}^{r^0}\left[\frac{(1-\mu_0)(r^0-P^0)}{nP^0\mu_0}\right]^{\frac{1}{n-1}} dP^0\right] = -\left(\left[\frac{r^0-P_0}{P_0}\right]^{\frac{1}{n-1}}\right)\frac{dP_0}{\partial \mu_0}$$

$$\text{But, } P_0 = \frac{(1-\mu_0)r^0}{1+n\mu_0-\mu_0}$$

$$\Rightarrow \frac{dP_0}{\partial \mu_0} = r^0\left[\frac{(1+n\mu_0-\mu_0)(-1)-(1-\mu_0)(n-1)}{(1+n\mu_0-\mu_0)^2}\right] = \frac{-(r^0 n)}{(1+n\mu_0-\mu_0)^2}$$

$$\text{Therefore, } \frac{\partial v}{\partial \mu_0} = -\left(\left[\frac{r^0-P_0}{P_0}\right]^{\frac{1}{n-1}}\right)\frac{dP_0}{\partial \mu_0} = -\left(\left[\frac{r^0-P_0}{P_0}\right]^{\frac{1}{n-1}}\right)\left(\frac{-(r^0 n)}{(1+n\mu_0-\mu_0)^2}\right)$$

$$\text{But } \frac{r^0-P_0}{P_0} = \frac{r^0}{P_0}-1 \text{ and } P_0 = \frac{(1-\mu_0)r^0}{1+n\mu_0-\mu_0},$$

$$\text{Thus } \left[\frac{r^0-P_0}{P_0}\right]^{\frac{1}{n-1}} = \left[\frac{n\mu_0}{1-\mu_0}\right]^{\frac{1}{n-1}}$$

$$\text{Therefore, } \frac{\partial v}{\partial \mu_0} = -\left[\frac{n\mu_0}{1-\mu_0}\right]^{\frac{1}{n-1}}\left(\frac{-(r^0 n)}{(1+n\mu_0-\mu_0)^2}\right) = \left[\frac{n\mu_0}{1-\mu_0}\right]^{\frac{1}{n-1}}\left(\frac{(r^0 n)}{(1+n\mu_0-\mu_0)^2}\right)$$

Also, $I = u.v$, hence $\frac{\partial I}{\partial \mu_0} = v\frac{\partial u}{\partial \mu_0} + u\frac{\partial v}{\partial \mu_0}$, therefore

$$\frac{\partial I}{\partial \mu_0} = \int_{P_0}^{r^0}\left[\frac{(1-\mu_0)(r^0-P^0)}{nP^0\mu_0}\right]^{\frac{1}{n-1}} dP^0\left((1/n)^{\frac{1}{n-1}}\left[-\left(\frac{1}{n-1}\right)\left(\frac{(1-\mu_0)^{\frac{1}{n-1}}}{(\mu_0)^{\frac{1}{n-1}+1}} + (1-\mu_0/\mu_0)^{\frac{1}{n-1}}\frac{1}{1-\mu_0}\right)\right]\right)$$

$$+\left[\frac{1-\mu_0}{n\mu_0}\right]^{1/n-1}\left[\frac{n\mu_0}{1-\mu_0}\right]^{\frac{1}{n-1}}\left(\frac{(r^0 n)}{(1+n\mu_0-\mu_0)^2}\right)$$

$$\Rightarrow \frac{\partial I}{\partial \mu_0} = \int_{P_0}^{r^0}\left[\frac{(1-\mu_0)(r^0-P^0)}{nP^0\mu_0}\right]^{\frac{1}{n-1}} dP^0\left((1/n)^{\frac{1}{n-1}}\left[-\left(\frac{1}{n-1}\right)\left(\frac{(1-\mu_0)^{\frac{1}{n-1}}}{(\mu_0)^{\frac{1}{n-1}+1}} + (1-\mu_0/\mu_0)^{\frac{1}{n-1}}\frac{1}{1-\mu_0}\right)\right]\right)$$

$$+\left(\frac{(r^0 n)}{(1+n\mu_0-\mu_0)^2}\right)$$

$$\text{Now, } \frac{\partial E(P^0)}{\partial \mu_0} = \frac{\partial P^0}{\partial \mu_0} + \frac{\partial I}{\partial \mu_0}$$

$$\Rightarrow \frac{\partial E(P^0)}{\partial \mu_0} = \frac{-(r^0 n)}{(1+n\mu_0-\mu_0)^2} +$$

$$\int_{P_0}^{r^0} \left[\frac{(1-\mu_0)(r^0-P^0)}{nP^0\mu_0}\right]^{\frac{1}{n-1}} dP^0 \left((1/n)^{\frac{1}{n-1}} \left[-\left(\frac{1}{n-1}\right)\left(\frac{(1-\mu_0)^{\frac{1}{n-1}}}{(\mu_0)^{\frac{1}{n-1}+1}} + (1-\mu_0/\mu_0)^{\frac{1}{n-1}} \frac{1}{1-\mu_0} \right)\right]\right)$$

$$+\left(\frac{(r^0 n)}{(1+n\mu_0-\mu_0)^2}\right)$$

$$\Rightarrow \frac{\partial E(P^0)}{\partial \mu_0} = \left((1/n)^{\frac{1}{n-1}} \left[-\left(\frac{1}{n-1}\right)\left(\frac{(1-\mu_0)^{\frac{1}{n-1}}}{(\mu_0)^{\frac{1}{n-1}+1}} + (1-\mu_0/\mu_0)^{\frac{1}{n-1}} \frac{1}{1-\mu_0} \right)\right]\right) \int_{P_0}^{r^0} \left[\frac{(1-\mu_0)(r^0-P^0)}{nP^0\mu_0}\right]^{\frac{1}{n-1}} dP^0$$

$$\Rightarrow \frac{\partial E(P^0)}{\partial \mu_0} < 0$$

Similarly we can prove it for the internet market.

Proposition 2: *Expected price of the searchers and the non-searchers in the overall market will decrease as the access to the internet will rise, if the fraction of the high intensity searchers in each market is the same.*

Proof:

$$\text{or, } F(P^0_{\min}) = 1-[1-F(P^0)]^n = 1-\left[\frac{(1-\mu_0)(r^0-P^0)}{nP^0\mu_0}\right]^{\frac{n}{n-1}}$$

$$f(P_{\min}^{0}) = -\left(\frac{n}{n-1}\right)\left(\frac{(1-\mu_0)}{n\mu_0}\right)^{\frac{n}{n-1}}\left(\frac{(r^0 - P^0)}{P^0}\right)^{\frac{n}{n-1}-1}\left(\frac{-r^0}{(P^0)^2}\right)$$

$$= \left(\frac{r^0}{(P^0)^2}\right)\frac{1}{n-1}\left(\frac{(1-\mu_0)}{n\mu_0}\right)^{\frac{n}{n-1}}\left(\frac{(r^0 - P^0)}{P^0}\right)^{\frac{1}{n-1}}$$

$$> E(P_{\min}^{0}) = \int_{P_0}^{r^0} \frac{r^0}{P^0(n-1)}\left(\frac{(1-\mu_0)}{\mu_0}\right)^{\frac{n}{n-1}}\left(\frac{(r^0 - P^0)}{nP^0}\right)^{\frac{1}{n-1}} dP^0$$

Similarly, $E(P_{\min}^{i}) = \int_{P_i}^{r^i} \frac{r^i}{P^i(n-1)}\left(\frac{(1-\mu_i)}{\mu_i}\right)^{\frac{n}{n-1}}\left(\frac{(r^i - P^i)}{nP^i}\right)^{\frac{1}{n-1}} dP^i$

Now, $E(P_{\min}^{overall}) = \frac{\lambda\mu_i}{\lambda\mu_i + (1-\lambda)\mu_0} E(P_{\min}^{i}) + \frac{(1-\lambda)\mu_0}{\lambda\mu_i + (1-\lambda)\mu_0} E(P_{\min}^{0})$

$$\Rightarrow \frac{\partial E(P_{\min}^{overall})}{\partial \lambda} = \frac{\mu_i\mu_0}{[\lambda\mu_i + (1-\lambda)\mu_0]^2} E(P_{\min}^{i}) - \frac{\mu_i\mu_0}{[\lambda\mu_i + (1-\lambda)\mu_0]^2} E(P_{\min}^{0})$$

$$= \frac{\mu_i\mu_0}{[\lambda\mu_i + (1-\lambda)\mu_0]^2}[E(P_{\min}^{i}) - E(P_{\min}^{0})]$$

where $\frac{\lambda\mu_i}{\lambda\mu_i + (1-\lambda)\mu_0}$ is the probability that the individual has access to the internet and also belongs to the group of high intensive searchers in the internet market. Likewise, $\frac{\lambda\mu_0}{\lambda\mu_i + (1-\lambda)\mu_0}$ refers to the probability that the individual is a high intensive searcher in the offline market.

Now, $E(P_{\min}^{i})$ and $E(P_{\min}^{0})$ involves an integral value which does not have a general solution since the value of the integral depends on the value of the number of sellers, n. However, for the general case, if $\mu_0 = \mu_i$ then $P_i < P_0$ and $r^0 > r^i$ by δ, such that $\frac{\partial E(P_{\min}^{overall})}{\partial \lambda} < 0$.

For n=2 we have:

$$E(P^0_{\min}) = \int_{P_0}^{r^0} \frac{r^0}{P^0(n-1)} \frac{(1-\mu_0)^2(r^0 - P^0)}{2\mu_0 P^0} dP^0 = \frac{(1-\mu_0)^2}{2\mu_0} \int_{P_0}^{r^0} \left(\frac{(r^0)^2}{(P^0)^2} - \frac{r^0}{P^0}\right) dP^0$$

$$= r^0 \left[\frac{(1-\mu_0)}{\mu_0} + \frac{(1-\mu_0)^2}{2.\mu_0} \ln \frac{(1-\mu_0)}{(1+\mu_0)} \right]$$

But, $r^0 = \dfrac{c_0}{1 - \left(\dfrac{(1-\mu_0)}{2.\mu_0}\right) . \ln \dfrac{(1+\mu_0)}{(1-\mu_0)}}$

Therefore, $E(P^0_{\min}) = c_0 \left[\dfrac{\dfrac{(1-\mu_0)}{\mu_0} + \dfrac{(1-\mu_0)^2}{2.\mu_0} \ln \dfrac{(1-\mu_0)}{(1+\mu_0)}}{1 - \left(\dfrac{(1-\mu_0)}{2.\mu_0}\right) . \ln \dfrac{(1+\mu_0)}{(1-\mu_0)}} \right]$

Similarly, $E(P^i_{\min}) = \delta c_0 \left[\dfrac{\dfrac{(1-\mu_i)}{\mu_i} + \dfrac{(1-\mu_i)^2}{2.\mu_i} \ln \dfrac{(1-\mu_i)}{(1+\mu_i)}}{1 - \left(\dfrac{(1-\mu_i)}{2.\mu_i}\right) . \ln \dfrac{(1+\mu_i)}{(1-\mu_i)}} \right]$

Suppose if $\dfrac{\partial E(P^{overall}_{\min})}{\partial \lambda} = \dfrac{\mu_i \mu_0}{[\lambda \mu_i + (1-\lambda)\mu_0]^2} [E(P^i_{\min}) - E(P^0_{\min})] < 0$ then,

$$\delta c_0 \left[\frac{\dfrac{(1-\mu_i)}{\mu_i} + \dfrac{(1-\mu_i)^2}{2.\mu_i} \ln \dfrac{(1-\mu_i)}{(1+\mu_i)}}{1 - \left(\dfrac{(1-\mu_i)}{2.\mu_i}\right) . \ln \dfrac{(1+\mu_i)}{(1-\mu_i)}} \right] < c_0 \left[\frac{\dfrac{(1-\mu_0)}{\mu_0} + \dfrac{(1-\mu_0)^2}{2.\mu_0} \ln \dfrac{(1-\mu_0)}{(1+\mu_0)}}{1 - \left(\dfrac{(1-\mu_0)}{2.\mu_0}\right) . \ln \dfrac{(1+\mu_0)}{(1-\mu_0)}} \right]$$

$$\Rightarrow \delta < \frac{\left[\dfrac{\dfrac{(1-\mu_0)}{\mu_0} + \dfrac{(1-\mu_0)^2}{2.\mu_0} \ln \dfrac{(1-\mu_0)}{(1+\mu_0)}}{1 - \left(\dfrac{(1-\mu_0)}{2.\mu_0}\right) . \ln \dfrac{(1+\mu_0)}{(1-\mu_0)}} \right]}{\left[\dfrac{\dfrac{(1-\mu_i)}{\mu_i} + \dfrac{(1-\mu_i)^2}{2.\mu_i} \ln \dfrac{(1-\mu_i)}{(1+\mu_i)}}{1 - \left(\dfrac{(1-\mu_i)}{2.\mu_i}\right) . \ln \dfrac{(1+\mu_i)}{(1-\mu_i)}} \right]}$$

Given below is the value of the ratio in the above bracket for different values of μ when **n=2**

Table A4
Sensitivity of Searcher's Search Cost to Different Values of μ

μ	Numerator	Denominator	Ratio
0.1	0.87283683	0.09698187	9
0.2	0.75627914	0.189069784	4
0.3	0.64817104	0.2778759	2.33
0.4	0.54678991	0.364526605	1.5
0.5	0.45069386	0.450693856	1
0.6	0.35860125	0.53790188	0.67
0.7	0.26927133	0.628299774	0.43
0.8	0.18133673	0.725346928	0.25
0.9	0.09293556	0.836420057	0.11

where $Numerator = \frac{(1-\mu)}{\mu} + \frac{(1-\mu)^2}{2.\mu}\ln\frac{(1-\mu)}{(1+\mu)}$ and

$$Denominator = 1 - \left(\frac{(1-\mu)}{2.\mu}\right).\ln\frac{(1+\mu)}{(1-\mu)}$$

From the value of the ratios provided in the table above, it is apparent that as the value of μ_s where s = i, 0 increases the ratio decreases. But we have already imposed the restriction that 0 < δ < 1. And we can see, from the table provided below, 0 < δ < 1 holds for all possible cases: (a) if $\mu_0 = \mu_i$ (b) if $\mu_0 > \mu_i$ (but with some restrictive values) and (c) if $\mu_0 < \mu_i$. Hence our supposition that $\frac{\partial E(P_{\min}^{overall})}{\partial \lambda} = [E(P_{\min}^{i}) - E(P_{\min}^{0})] < 0$ is true irrespective of the relationship between μ_0 and μ_i. This is however not true, if we want to claim that $\frac{\partial E(P_{\min}^{overall})}{\partial \lambda} > 0$, since in this case the restriction on δ will cease to hold.

For example if $\mu_0 = \mu_i$, then for $\frac{\partial E(P_{\min}^{overall})}{\partial \lambda} > 0$ to hold δ has to be greater than 1, which is not feasible given our assumption.

However, only if, $1 > \delta > \dfrac{\left[\dfrac{\dfrac{(1-\mu_0)}{\mu_0}+\dfrac{(1-\mu_0)^2}{2.\mu_0}\ln\dfrac{(1-\mu_0)}{(1+\mu_0)}}{1-\left(\dfrac{(1-\mu_0)}{2.\mu_0}\right).\ln\dfrac{(1+\mu_0)}{(1-\mu_0)}}\right]}{\left[\dfrac{\dfrac{(1-\mu_i)}{\mu_i}+\dfrac{(1-\mu_i)^2}{2.\mu_i}\ln\dfrac{(1-\mu_i)}{(1+\mu_i)}}{1-\left(\dfrac{(1-\mu_i)}{2.\mu_i}\right).\ln\dfrac{(1+\mu_i)}{(1-\mu_i)}}\right]}$ and $\mu_0 > \mu_i$, do we have $\dfrac{\partial E(P_{\min}^{overall})}{\partial\lambda} > 0$. Or in other words only in a situation where the fraction of searchers in the offline market is greater than that of their counterpart in the internet market, may generate a $1 > \delta > \dfrac{\left[\dfrac{\dfrac{(1-\mu_0)}{\mu_0}+\dfrac{(1-\mu_0)^2}{2.\mu_0}\ln\dfrac{(1-\mu_0)}{(1+\mu_0)}}{1-\left(\dfrac{(1-\mu_0)}{2.\mu_0}\right).\ln\dfrac{(1+\mu_0)}{(1-\mu_0)}}\right]}{\left[\dfrac{\dfrac{(1-\mu_i)}{\mu_i}+\dfrac{(1-\mu_i)^2}{2.\mu_i}\ln\dfrac{(1-\mu_i)}{(1+\mu_i)}}{1-\left(\dfrac{(1-\mu_i)}{2.\mu_i}\right).\ln\dfrac{(1+\mu_i)}{(1-\mu_i)}}\right]}$ such that the expected price to be paid by the searchers in the overall market will rise.

Hence, dominance of high intensive searchers in the offline market may dominate the rise in the internet usage effect, causing the expected price to be paid by the high intensive searcher in the overall market to rise. Though theoretically this may seem intriguing, yet, the existence of such a condition in the real world must be seen with much skepticism. Our thinking is that, with the internet, the fraction of searchers in the two market maybe at the least equal with a much higher probability that the fraction of searchers in the internet market is higher than that of their offline counterpart, $\mu_0 < \mu_i$.

For non-searchers:

$$E(P_{non-searchers}^{overall}) = \frac{\lambda(1-\mu_i)}{\lambda(1-\mu_i)+(1-\lambda)(1-\mu_0)}E(P^i) + \frac{(1-\lambda)(1-\mu_0)}{\lambda(1-\mu_i)+(1-\lambda)(1-\mu_0)}E(P^0)$$

$$\frac{\partial E(P^{overall}_{non-searchers})}{\partial \lambda} = \frac{(1-\mu_i)(1-\mu_0)}{[\lambda(1-\mu_i)+(1-\lambda)(1-\mu_0)]^2}[E(P^i)-E(P^0)]$$

where $\dfrac{\lambda(1-\mu_i)}{\lambda(1-\mu_i)+(1-\lambda)(1-\mu_0)}$ is the probability that the individual has access to the internet but belongs to the group of low intensive searchers in the internet market. Similarly, $\dfrac{(1-\lambda)(1-\mu_0)}{\lambda(1-\mu_i)+(1-\lambda)(1-\mu_0)}$ refers to the probability that the individual is a high intensive searcher in the offline market.

If $\mu_0 = \mu_i$, then $E(P^i) < E(P^0) \Rightarrow \dfrac{\partial E(P^{overall}_{non-searchers})}{\partial \lambda} < 0$. As before the $E(P^i_{\min})$ and $E(P^0_{\min})$ involves an integral value which does not have a general solution since the value of the integral depends on the value of the numbers of sellers, n and that $\dfrac{\partial E(P^{overall}_{non-searchers})}{\partial \lambda} < 0$ is easily visualized only in the case when $\mu_0 = \mu_i$. However for a more analytical proof of our claim we consider the simple duopoly case.

We know that,

$$E(P^0) = P_0 + \int_{P_0}^{r^0}\left[\frac{(1-\mu_0)(r^0-P^0)}{nP^0\mu_0}\right]^{\frac{1}{n-1}} dP^0$$

For n = 2,

$$E(P^0) = P_0 + \int_{P_0}^{r^0}\left[\frac{(1-\mu_0)(r^0-P^0)}{2P^0\mu_0}\right] dP^0 = P_0 + \frac{(1-\mu_0)}{2\mu_0}\int_{P_0}^{r^0}\frac{(r^0-P^0)}{P^0} dP^0$$

Using the fact that $P_0 = \dfrac{(1-\mu_0)r^0}{1+n\mu_0-\mu_0}$ and by some mathematical manipulations we get:

$$E(P^0) = -(\frac{(1-\mu_0)}{2\mu_0})\frac{c_0}{1-\left(\frac{(1-\mu_0)}{2\mu_0}\right)\ln\left(\frac{(1+\mu_0)}{(1-\mu_0)}\right)}\ln\left(\frac{(1-\mu_0)}{(1+\mu_0)}\right)$$

Similarly,

$$E(P^i) = -(\frac{(1-\mu_i)}{2\mu_i})\frac{\delta c_0}{1-\left(\frac{(1-\mu_i)}{2\mu_i}\right)\ln\left(\frac{(1+\mu_i)}{(1-\mu_i)}\right)}\ln\left(\frac{(1-\mu_i)}{(1+\mu_i)}\right)$$

For $\frac{\partial E(P^{overall}_{non-searchers})}{\partial \lambda}$ to be negative, we require $E(P^i) < E(P^0)$ and 0<1<δ to hold.

Now $E(P^i) < E(P^0)$ implies:

$$-(\frac{(1-\mu_i)}{2\mu_i})\frac{\delta c_0}{1-\left(\frac{(1-\mu_i)}{2\mu_i}\right)\ln\left(\frac{(1+\mu_i)}{(1-\mu_i)}\right)}\ln\left(\frac{(1-\mu_i)}{(1+\mu_i)}\right) <$$

$$-(\frac{(1-\mu_0)}{2\mu_0})\frac{c_0}{1-\left(\frac{(1-\mu_0)}{2\mu_0}\right)\ln\left(\frac{(1+\mu_0)}{(1-\mu_0)}\right)}\ln\left(\frac{(1-\mu_0)}{(1+\mu_0)}\right)$$

$$\Rightarrow \delta < \frac{(\frac{(1-\mu_0)}{2\mu_0})\frac{c_0}{1-\left(\frac{(1-\mu_0)}{2\mu_0}\right)\ln\left(\frac{(1+\mu_0)}{(1-\mu_0)}\right)}\ln\left(\frac{(1-\mu_0)}{(1+\mu_0)}\right)}{(\frac{(1-\mu_i)}{2\mu_i})\frac{\delta c_0}{1-\left(\frac{(1-\mu_i)}{2\mu_i}\right)\ln\left(\frac{(1+\mu_i)}{(1-\mu_i)}\right)}\ln\left(\frac{(1-\mu_i)}{(1+\mu_i)}\right)}$$

The table below gives the value of the ratio in the bracket for different values of μ when n = 2:

Table A5
Sensitivity of Non-searcher's Search Cost to Different Values of μ

μ	Numerator (N)	Denominator (D)	Ratio = -N/D
0.1	-0.90301813	0.09698187	9.311
0.2	-0.810930216	0.189069784	4.289
0.3	-0.72221241	0.27778759	2.6
0.4	-0.635473395	0.364526605	1.74
0.5	-0.549306144	0.450693856	1.218
0.6	-0.46209812	0.53790188	0.859
0.7	-0.371700226	0.628299774	0.591
0.8	-0.274653072	0.725346928	0.378
0.9	-0.163579943	0.836420057	0.195

where $Numerator = \frac{(1-\mu)}{2.\mu} \ln \frac{(1-\mu)}{(1+\mu)}$ and

$$Deno\min ator = 1 - \left(\frac{(1-\mu)}{2.\mu} \right) . \ln \frac{(1+\mu)}{(1-\mu)}$$

Thus, we see as before, as the value of λ increases the ratio starts to decrease. In such a situation our restriction on δ will hold similarly in all of the three cases as mentioned before. Also if we want to argue that $\frac{\partial E(P_{non-searchers}^{overall})}{\partial \lambda} > 0$ then our assumption about δ ceases to hold. Hence it can be argued strongly that both $\frac{\partial E(P_{non-searchers}^{overall})}{\partial \lambda} < 0$ and $\frac{\partial E(P_{\min}^{overall})}{\partial \lambda} < 0$. Here too we have a situation similar to that of the high intensive searchers.

If $1 > \delta > \frac{\left[\dfrac{\frac{(1-\mu_0)}{\mu_0} + \frac{(1-\mu_0)^2}{2.\mu_0} \ln \frac{(1-\mu_0)}{(1+\mu_0)}}{1 - \left(\frac{(1-\mu_0)}{2.\mu_0} \right) . \ln \frac{(1+\mu_0)}{(1-\mu_0)}} \right]}{\left[\dfrac{\frac{(1-\mu_i)}{\mu_i} + \frac{(1-\mu_i)^2}{2.\mu_i} \ln \frac{(1-\mu_i)}{(1+\mu_i)}}{1 - \left(\frac{(1-\mu_i)}{2.\mu_i} \right) . \ln \frac{(1+\mu_i)}{(1-\mu_i)}} \right]}$ and $\mu_0 > \mu_i$, then the expected price to be paid by the low intensive searcher in the overall market will rise as the access to the internet rises, contradicting the popular belief that greater access to information intensifies competition, and thereby lower the average price.

Corollary 2.1: *For the duopoly, the expected price to be paid by the high intensive searcher in the internet market will be lower than their offline counterpart.*

Proof:

This follows directly from the fact that $\frac{\partial E(P_{\min}^{overall})}{\partial \lambda} < 0$ and

(a) $\mu_0 < \mu_i$, with $\delta < \dfrac{\left[\dfrac{\dfrac{(1-\mu_0)}{\mu_0}+\dfrac{(1-\mu_0)^2}{2.\mu_0}\ln\dfrac{(1-\mu_0)}{(1+\mu_0)}}{1-\left(\dfrac{(1-\mu_0)}{2.\mu_0}\right).\ln\dfrac{(1+\mu_0)}{(1-\mu_0)}}\right]}{\left[\dfrac{\dfrac{(1-\mu_i)}{\mu_i}+\dfrac{(1-\mu_i)^2}{2.\mu_i}\ln\dfrac{(1-\mu_i)}{(1+\mu_i)}}{1-\left(\dfrac{(1-\mu_i)}{2.\mu_i}\right).\ln\dfrac{(1+\mu_i)}{(1-\mu_i)}}\right]}$

(b) $\mu_0 = \mu_i$, with $\delta < \dfrac{\left[\dfrac{\dfrac{(1-\mu_0)}{\mu_0}+\dfrac{(1-\mu_0)^2}{2.\mu_0}\ln\dfrac{(1-\mu_0)}{(1+\mu_0)}}{1-\left(\dfrac{(1-\mu_0)}{2.\mu_0}\right).\ln\dfrac{(1+\mu_0)}{(1-\mu_0)}}\right]}{\left[\dfrac{\dfrac{(1-\mu_i)}{\mu_i}+\dfrac{(1-\mu_i)^2}{2.\mu_i}\ln\dfrac{(1-\mu_i)}{(1+\mu_i)}}{1-\left(\dfrac{(1-\mu_i)}{2.\mu_i}\right).\ln\dfrac{(1+\mu_i)}{(1-\mu_i)}}\right]}$

(c) $\mu_0 > \mu_i$, with $\delta < \dfrac{\left[\dfrac{\dfrac{(1-\mu_0)}{\mu_0}+\dfrac{(1-\mu_0)^2}{2.\mu_0}\ln\dfrac{(1-\mu_0)}{(1+\mu_0)}}{1-\left(\dfrac{(1-\mu_0)}{2.\mu_0}\right).\ln\dfrac{(1+\mu_0)}{(1-\mu_0)}}\right]}{\left[\dfrac{\dfrac{(1-\mu_i)}{\mu_i}+\dfrac{(1-\mu_i)^2}{2.\mu_i}\ln\dfrac{(1-\mu_i)}{(1+\mu_i)}}{1-\left(\dfrac{(1-\mu_i)}{2.\mu_i}\right).\ln\dfrac{(1+\mu_i)}{(1-\mu_i)}}\right]}$

Corollary 2.2: *For the duopoly, the expected price to be paid by the low intensive searcher in the internet market will be lower than their offline counterpart.*

Proof:

Follows as a consequence of $\dfrac{\partial E(P^{overall}_{non-searchers})}{\partial \lambda} < 0$ and

(a) $\mu_0 < \mu_i$, with $\delta < \dfrac{(\dfrac{(1-\mu_0)}{2\mu_0})\dfrac{c_0}{1-\left(\dfrac{(1-\mu_0)}{2\mu_0}\right)\ln\left(\dfrac{(1+\mu_0)}{(1-\mu_0)}\right)}\ln\left(\dfrac{(1-\mu_0)}{(1+\mu_0)}\right)}{(\dfrac{(1-\mu_i)}{2\mu_i})\dfrac{\delta c_0}{1-\left(\dfrac{(1-\mu_i)}{2\mu_i}\right)\ln\left(\dfrac{(1+\mu_i)}{(1-\mu_i)}\right)}\ln\left(\dfrac{(1-\mu_i)}{(1+\mu_i)}\right)}$

(b) $\mu_0 = \mu_i$, with $\delta < \dfrac{(\dfrac{(1-\mu_0)}{2\mu_0})\dfrac{c_0}{1-\left(\dfrac{(1-\mu_0)}{2\mu_0}\right)\ln\left(\dfrac{(1+\mu_0)}{(1-\mu_0)}\right)}\ln\left(\dfrac{(1-\mu_0)}{(1+\mu_0)}\right)}{(\dfrac{(1-\mu_i)}{2\mu_i})\dfrac{\delta c_0}{1-\left(\dfrac{(1-\mu_i)}{2\mu_i}\right)\ln\left(\dfrac{(1+\mu_i)}{(1-\mu_i)}\right)}\ln\left(\dfrac{(1-\mu_i)}{(1+\mu_i)}\right)}$

(c) $\mu_0 > \mu_i$, with $\delta < \dfrac{(\dfrac{(1-\mu_0)}{2\mu_0})\dfrac{c_0}{1-\left(\dfrac{(1-\mu_0)}{2\mu_0}\right)\ln\left(\dfrac{(1+\mu_0)}{(1-\mu_0)}\right)}\ln\left(\dfrac{(1-\mu_0)}{(1+\mu_0)}\right)}{(\dfrac{(1-\mu_i)}{2\mu_i})\dfrac{\delta c_0}{1-\left(\dfrac{(1-\mu_i)}{2\mu_i}\right)\ln\left(\dfrac{(1+\mu_i)}{(1-\mu_i)}\right)}\ln\left(\dfrac{(1-\mu_i)}{(1+\mu_i)}\right)}$

Proposition 3: *Prices are dispersed over a wider range of prices in the overall market in comparison to an individual market, online and offline.*

Proof: We have seen earlier that $M(P) = \lambda.G(P^i) + (1-\lambda).F(P^0)$. Now the expected price in the overall market, consisting of both the online and the offline market is given by:

$$\text{E}(\ E[M(P)] = \lambda.G(P^i) + (1-\lambda).F(P^0)$$

Now the expected price in the overall market, consisting of both the online and offline markets is given by

$E[M(P)] = \lambda E[G(P^i)] + (1-\lambda)E[F(P^o)]$

or, $E[M(P)] = \lambda\{\mu_i E(P^i_{\min}) + (1-\mu_i)E(P^i)\} + (1-\lambda)\{\mu_o E(P^o_{\min}) + (1-\mu_o)E(P^o)\}$

or, $E[M(P)] = \lambda B(\mu_i) + (1-\lambda)B(\mu_o) = \lambda(B(\mu_i) - B(\mu_o)) + B(\mu_o)$

where, $B(\mu_i) = \mu_i E(P^i_{\min}) + (1-\mu_i)E(P^i)$

and, $B(\mu_o) = \mu_o E(P^o_{\min}) + (1-\mu_o)E(P^o)$

Also,

$$E[M(P^2)] = \lambda\{\mu_i E((P^i_{\min})^2) + (1-\mu_i)E((P^i)^2)\} + (1-\lambda)\{\mu_o E((P^o_{\min})^2) + (1-\mu_0)E((P^o)^2)\}$$

or, $E[M(P^2)] = \lambda\{\mu_i H_1(\mu_i) + (1-\mu_i)H_2(\mu_i)\} + (1-\lambda)\{\mu_0 H_3(\mu_0) + (1-\mu_0)H_4(\mu_o)\}$

where, $H_1(\mu_i) = E((P^i_{\min})^2)$,

$H_2(\mu_i) = E((P^i)^2)$,

$H_3(\mu_0) = E((P^o_{\min})^2$

and, $H_4(\mu_o) = E((P^o)^2)$

Now, $\text{var}[M(P)] = E[M(P^2)] - [E[M(P)]]^2$

Plugging in the expressions for $E[M(P^2)]$ and $[E[M(P)]]^2$ we get:

$$\frac{\partial \text{var}[M(P)]}{\partial \lambda} = \mu_i H_1(\mu_i) + H_2(\mu_i) - \mu_i H_2(\mu_i) - \mu_0 H_3(\mu_0) - H_4(\mu_0)$$
$$+ \mu_o H_4(\mu_0) - 2\lambda[B(\mu_i) - B(\mu_0)]^2 - 2[B(\mu_i) - B(\mu_0)]B(\mu_0)$$

Simplification of the above expression yields that $\frac{\partial \text{var}[M(P)]}{\partial \lambda} \gtrless 0$

Further, $\frac{\partial^2 \text{var}[M(P)]}{\partial \lambda^2} = -2[B(\mu_i) - B(\mu_0)]^2 < 0$

It is straightforward to realize that in the overall market the lowest price is P_i and the higher price is r^o, reservation price in the offline market. It can be concluded, then that the distribution of prices in the overall market will be more dispersed as compared the distribution in any individual market.

VITA

Name: Anirban Sengupta

Address: Analysis Group, Two Embarcadero Center, Suite 1750, San Francisco, CA-94111

E-mail: anirban.sengupta@gmail.com

EDUCATION:

Texas A&M University, PhD in Economics, August 2007.

Jawaharlal Nehru University, New Delhi, India, Master of Arts in Economics, July 2000.

St. Xavier's College, Kolkata, India, Bachelor of Science, July 1998.

ACADEMIC GRANTS, AWARDS AND DISTINCTIONS:

Best Session Paper, Student Research Week, Texas A&M University, March 2007.

Networks, Electronic Commerce and Telecommunications Institute (NET institute) Research Grant, Summer 2006.

Charles Maurice Fellowship for Excellence in Graduate Studies, Department of Economics, Texas A&M University, 2004-2005.

Graduate Teaching Assistant Excellence Award, Department of Economics, Texas A&M University, 2004-2005.